# 化工方向多尺度特征融合文献 - 按算法类型分类

## Deep Learning (Other)

### 1. Multiscale Feature Fusion-Based Object Detection Algorithm

摘要: The RetinaNet and Libra RetinaNet object detectors based on deep learning employ feature pyramid networks to fuse multiscale features. However, insufficient feature fusion is problematic in these detectors. In this paper, a multiscale feature fusion algorithm is proposed. The proposed algorithm is extended based on Libra RetinaNet. Two independent feature fusion modules are constructed by establishing two bottom-up paths, and the results generated by the two modules are fused with the original predicted features to improve the accuracy of the detector. The multiscale feature fusion module and Libra RetinaNet are combined to build a target detector and conduct experiments on different datasets. Experimental results demonstrate that the average accuracy of the added module detector on PASCAL VOC and MSCOCO datasets is improved by 2.2 and 1.3 percentage, respectively, compared to the Libra RetinaNet detector.

### 2. Fault Diagnosis for Rolling Bearings Based on Multiscale Feature Fusion

摘要: Deep learning, due to its excellent feature-adaptive capture ability, has been widely utilized in the fault diagnosis field. However, there are two common problems in deep-learning-based fault diagnosis methods: (1) many researchers attempt to deepen the layers of deep learning models for higher diagnostic accuracy, but degradation problems of deep learning models often occur; and (2) the use of multiscale features can easily be ignored, which makes the extracted data features lack diversity. To deal with these problems, a novel multiscale feature fusion deep residual network is proposed in this paper for the fault diagnosis of rolling bearings, one which contains multiple multiscale feature fusion blocks and a multiscale pooling layer. The multiple multiscale feature fusion block is designed to automatically extract the multiscale features from raw signals, and further compress them for higher dimensional feature mapping. The multiscale pooling layer is constructed to fuse the extracted multiscale feature mapping. Two famous rolling bearing datasets are adopted to evaluate the diagnostic performance of the proposed model. The comparison results show that the diagnostic performance of the proposed model is superior to not only several popular models, but also other advanced methods in the literature.

### 3. Multiscale feature fusion for surveillance video diagnosis

摘要: Recently, surveillance video diagnosis has attracted increasing interest for generating real-time alarms related to camera failure in video surveillance systems. The existing surveillance video diagnosis methods do not have sufficient ability to detect multiple types of anomalies. Therefore, this paper proposes a surveillance video diagnosis method based on deep learning to detect multiple types of anomalies. A multiscale feature fusion residual network is designed to detect and classify camera anomalies. The experimental results show that the classification accuracy of the proposed method is more than 98%. (C)& nbsp;2022 Elsevier B.V. All rights reserved.

### 4. nMultiscale Feature Fusion-Based Object Detection Algorithm

摘要: The RetinaNet and Libra RetinaNet object detectors based on deep learning employ feature pyramid networks to fuse multiscale features. However, insufficient feature fusion is problematic in these detectors. In this paper, a multiscale feature fusion algorithm is proposed. The proposed algorithm is extended based on Libra RetinaNet. Two independent feature fusion modules are constructed by establishing two bottom-up paths, and the results generated by the two modules are fused with the original predicted features to improve the accuracy of the detector. The multiscale feature fusion module and Libra RetinaNet are combined to build a target detector and conduct experiments on different datasets. Experimental results demonstrate that the average accuracy of the added module detector on PASCAL VOC and MSCOCO datasets is improved by 2. 2 and 1. 3 percentage, respectively, compared to the Libra RetinaNet detector.

### 5. Improving Geological Remote Sensing Interpretation Via a Contextually

摘要: Geological remote sensing interpretation plays a pivotal role in the field of regional geological mapping, encompassing the analysis of rock, soil, and water features. However, these geological elements can be obscured by the surrounding geographical environment and can undergo modifications caused by geological activities. The former hinders the effectiveness of satellite remote sensing data, resulting in the invisibility of element features, while the latter leads to the complex distribution of element features and significant spatial variations of geological elements. Consequently, existing deep learning-based models for interpreting geological elements often exhibit limited accuracy. To address these issues, this study proposes the contextually enhanced multiscale feature fusion network for the efficient interpretation of geological elements. First, the context enhancement module is employed to extract abundant feature information and reinforce contextual features, aiming to capture essential features and strengthen their interconnections. Second, the multiscale feature fusion module incorporates the SimAM attention mechanism to adaptively learn features from different channels, emphasizing the feature information that contributes to interpretation results and maximizing the comprehensive and crucial feature information for each element. Extensive experiments demonstrate the superior performance of both the context enhancement module and the multiscale feature fusion module compared to several representative deep learning networks in terms of overall interpretation accuracy on two datasets. The model demonstrated improvements in oPA and mIoU of 2.4% and 2.8%, respectively, on the Landsat 8 dataset, and 3.5% and 3.2%, respectively, on the Sentinel-2 dataset.

### 6. Progressive structure network-based multiscale feature fusion for object

摘要: Deep learning-based target detection techniques have already made a wide-range impact on our daily life. Currently, a feature pyramid is a widely utilized technique for multiscale target detection, the effectiveness of the technique has already been proved. Nevertheless, in the pyramid structure, problems, such as multiscale feature alignment, model turmoil after fusion, feature redundancy, and no-local feature fusion, exist. In this paper, we propose a novel progressive structure network to solve the aforementioned problems. The proposed structure contains three modules: multiscale feature alignment fusion, different scale channels & spatial location adaptive weighted fusion, and multiscale global and local feature fusion. The proposed structure is capable of fusing information from different feature layers more effectively. Subsequently, the semantic gaps among different scales can be reduced. Furthermore, the proposed structure can maintain the stability of the detection network and its performance has been proved by comparing with other state-of-art feature fusion method. The proposed progressive network structure has also been applied to actual target detection tasks and the practical application effectiveness of our method has been verified.

### 7. Dual discriminator residual generation adversarial network with

摘要: Generative adversarial networks (GANs), a class of unsupervised generative models based on deep learning, have shown promising results in generating realistic and high-quality images without modeling the data distribution. However, standard GANs are often difficult to train and often suffer from gradient disappearance, gradient explosion, or mode collapse, which can limit the overall performance of the model. To address the mode collapse problem and improve the variety of the generated images, this paper proposes a dual discriminator structure. Furthermore, this paper improves the generator and discriminator model and proposes a generator based on residual network and multiscale feature fusion and a discriminator based on multiscale feature fusion, which effectively solves the problem of gradient disappearance and gradient explosion that occurs in deep networks while improving the quality of generated images. The proposed approach is applied to various datasets, including MNIST, LSUN, and CelebA. The training results reveal that the stability and quality of the generated images are high, achieving satisfying FID and IS values.

### 8. Multiscale feature fusion network for 3D head MRI image registration

摘要: BackgroundImage registration technology has become an important medical image preprocessing step with the wide application of computer-aided diagnosis technology in various medical image analysis tasks. PurposeWe propose a multiscale feature fusion registration based on deep learning to achieve the accurate registration and fusion of head magnetic resonance imaging (MRI) and solve the problem that general registration methods cannot handle the complex spatial information and position information of head MRI. MethodsOur proposed multiscale feature fusion registration network consists of three sequentially trained modules. The first is an affine registration module that implements affine transformation; the second is to realize non-rigid transformation, a deformable registration module composed of top-down and bottom-up feature fusion subnetworks in parallel; and the third is a deformable registration module that also realizes non-rigid transformation and is composed of two feature fusion subnetworks in series. The network decomposes the deformation field of large displacement into multiple deformation fields of small displacement by multiscale registration and registration, which reduces the difficulty of registration. Moreover, multiscale information in head MRI is learned in a targeted manner, which improves the registration accuracy, by connecting the two feature fusion subnetworks. ResultsWe used 29 3D head MRIs for training and seven volumes for testing and calculated the values of the registration evaluation metrics for the new algorithm to register anterior and posterior lateral pterygoid muscles. The Dice similarity coefficient was 0.745 +/- 0.021, the Hausdorff distance was 3.441 +/- 0.935 mm, the Average surface distance was 0.738 +/- 0.098 mm, and the Standard deviation of the Jacobian matrix was 0.425 +/- 0.043. Our new algorithm achieved a higher registration accuracy compared with state-of-the-art registration methods. ConclusionsOur proposed multiscale feature fusion registration network can realize end-to-end deformable registration of 3D head MRI, which can effectively cope with the characteristics of large deformation displacement and the rich details of head images and provide reliable technical support for the diagnosis and analysis of head diseases.

### 9. A Multiscale Spatiotemporal Fusion Network Based on an Attention

摘要: Spatiotemporal fusion is an effective and cost-effective method to obtain both high temporal resolution and high spatial resolution images. However, existing methods do not sufficiently extract the deeper features of the image, resulting in fused images which do not recover good topographic detail and poor fusion quality. In order to obtain higher quality spatiotemporal fusion images, a novel spatiotemporal fusion method based on deep learning is proposed in this paper. The method combines an attention mechanism and a multiscale feature fusion network to design a network that more scientifically explores deeper features of the image for different input image characteristics. Specifically, a multiscale feature fusion module is introduced into the spatiotemporal fusion task and combined with an efficient spatial-channel attention module to improve the capture of spatial and channel information while obtaining more effective information. In addition, we design a new edge loss function and incorporate it into the compound loss function, which helps to generate fused images with richer edge information. In terms of both index performance and image details, our proposed model has excellent results on both datasets compared with the current mainstream spatiotemporal fusion methods.

### 10. Self-Attention Metric Learning Based on Multiscale Feature Fusion for

摘要: The intelligent diagnosis model based on deep learning method can effectively and accurately diagnose the health state of bearings widely used in various mechanical equipment. However, in engineering practice, it is difficult to obtain sufficient labeled fault data, which would reduce the diagnosis performance. In order to solve the above problem, this article proposes a self-attention metric learning based on multiscale feature fusion to classify the bearing fault with few shots. The proposed few-shot intelligent diagnosis model mainly contains the feature extraction module and metric learning module. First, the vibration data need to be combined in pairs to form sample pairs, which are input into the feature extraction module based on multiscale feature fusion (MSFF). This input approach can increase the learning tasks so as to alleviate the problem of insufficient training samples. Then, the extracted features are concatenated and input into the metric learning module whose learning ability is improved by adding the self-attention (SA) network, and the input features are eventually converted into the similarity. Finally, the fault type can be identified according to the maximum similarity value between the test sample and samples with different labels. Three experiment cases are conducted to validate the performance of proposed intelligent diagnosis model. The experimental results show that the proposed model can accurately classify the bearing fault with few-shot case. Additionally, the comparison experiments with the traditional models have also validated the advantage of the proposed intelligent diagnosis model for few-shot case.

### 11. A Photovoltaic Hot-Spot Fault Detection Network for Aerial Images Based

摘要: The number of samples is one of the key factors affecting the performance of deep learning-based detection networks. Aiming at the problem that the detection network is difficult to accurately detect the hot-spot fault targets under the condition of small samples, a photovoltaic hot-spot fault detection network based on progressive transfer learning and multiscale feature fusion is proposed. First, a large number of artificial hot-spot samples are generated through the artificial model, and the mixed dataset containing real and artificial samples is constructed to improve the data diversity. On this basis, a pre-trained model based on artificial samples is established to learn the shallow features of hot-spot faults. Then, to fuse the multiscale features and improve feature aggregation ability of detection network, a novel feature pyramid structure based on reparameterized generalized and multiscale feature fusion (RepG-MSFF) is designed. Moreover, to balance the detection accuracy and speed, the spatial and channel reconstruction convolution (SCConv) is utilized to replace conventional convolution in the backbone network. Finally, to further accurately locate hot-spot targets, an adaptive threshold focal loss (TFL) function is introduced. The experimental results indicate that, in three different scenarios datasets, the detection accuracy can reach 87.9%, 88.6%, and 87.7%, respectively, which is higher than that of other nine detection algorithms.

### 12. Multiscale Feature Interactive Network for Multifocus Image Fusion

摘要: In deep learning (DL)-based multifocus image fusion, effective multiscale feature learning is a key issue to promote fusion performance. In this article, we propose a novel DL model named multiscale feature interactive network (MSFIN), which can segment the source images into focused and defocused regions accurately by sufficient interaction of multiscale features from layers of different depths in the network for multifocus image fusion. Specifically, based on the popular encoder-decoder framework, two functional modules, namely, multiscale feature fusion (MSFF) and coordinate attention upsample (CAU), are designed for interactive multiscale feature learning. Moreover, the weighted binary cross-entropy (WBCE) loss and the multilevel supervision (MLS) strategy are introduced to train the network more effectively. Qualitative and quantitative comparisons with 19 representative multifocus image fusion methods demonstrate that the proposed method can achieve state-of-the-art performance. The code of our method is available at

### 13. A novel hierarchical structural pruning-multiscale feature fusion

摘要: Residual learning is a commonly used method in the intelligent fault diagnosis (IFD) field. The existing residual networks usually have a large model volume, whose training process puts for-ward high requirements for both time consumption and the computing resource configuration. This paper proposes a novel hierarchical structural pruning-multiscale feature fusion residual network (HSP-MFFRN) for IFD. The multiple multi-scale feature extraction modules and feature fusion modules are designed in the proposed HSP-MFFRN to extract, fuse and compress the multi -scale features without changing the size of the convolutional filter. In addition, hierarchical structural pruning is implemented for HSP-MFFRN to delete redundant channels. Two experi-mental study cases are conducted to verify model diagnosis performance, including a rolling bearing IFD case and an aerostat capsule IFD case. The experimental results show that the pro-posed model can effectively compress the learnable parameters and model volume after hierar-chical structural pruning, and can still achieve superior diagnostic performance compared with other deep learning methods, residual learning methods, and other advanced methods reported in the literature.

### 14. TAF2-Net: Triple-Branch Attentive Feature Fusion Network for Ultrasonic

摘要: Automatic defect detection is a critical task in the industrial production process. At present, many detection methods based on deep learning have been successfully applied in industrial defect detection systems. However, due to the complexity of the specimen defects, the data collected in practice are imbalanced and have intraclass differences and interclass similarities. These problems may reduce the accuracy of defect detection. This article proposes an ultrasonic defect detection system, which uses a novel triple-branch attentive feature fusion network (TAF2-Net) architecture to detect and classify the defects of specimens. We introduce the multiscale feature extraction (MSFE) module to extract multiscale features from the 1-D ultrasonic echo signal and 2-D time & x2013;frequency diagram in two branches and design the deep multimodal feature fusion (DMF2) module to enhance the features at different abstract levels. The fused features are unified through a sequential feature pyramid network (SFPN) and fed into gated recurrent unit (GRU) to extract global sequential features in the third branch. Finally, the features extracted from the multiscale feature extract-1-D (MSFE-1-D) module, multiscale feature extract-2-D (MSFE-2-D) module, and GRU are fused to obtain the final prediction results. Experimental results show that the TAF2-Net is superior to the most advanced method, with the accuracy of 96.54 & x0025;. Meanwhile, to meet the requirement of hardware systems that lack of computing resources, we also propose a lightweight TAF2-small Net that can still achieve a classification accuracy of 93.6 & x0025;.

### 15. Denoising-Based Multiscale Feature Fusion for Remote Sensing Image

摘要: With the benefits from deep learning technology, generating captions for remote sensing images has become achievable, and great progress has been made in this field in the recent years. However, a large-scale variation of remote sensing images, which would lead to errors or omissions in feature extraction, still limits the further improvement of caption quality. To address this problem, we propose a denoising-based multi-scale feature fusion (DMSFF) mechanism for remote sensing image captioning in this letter. The proposed DMSFF mechanism aggregates multiscale features with the denoising operation at the stage of visual feature extraction. It can help the encoder-decoder framework, which is widely used in image captioning, to obtain the denoising multiscale feature representation. In experiments, we apply the proposed DMSFF in the encoder-decoder framework and perform the comparative experiments on two public remote sensing image captioning data sets including UC Merced (UCM)-captions and Sydney-captions. The experimental results demonstrate the effectiveness of our method.

### 16. DSMFFNet: Depthwise separable multiscale feature fusion network for

摘要: Bridge detection methods based on deep learning have many parameters, complex calculations, and serious errors and missed detections for multiscale bridges. To solve the above problems, a depth-wise separable multiscale feature fusion network (DSMFFNet) is proposed for efficient and accurate bridge detection in very high resolution satellite images (VHR). First, depth-wise separable convolution was used to build a backbone feature extraction network to extract the bridge features. Second, to better match bridges of different scales, multiscale receptive fields were obtained by multibranch parallel dilated convolution at the last layer of the feature map. Then, to make full use of the details and semantic information of the bridges at different depths, the three effective feature layers of the bridges at different levels are fused by a multiscale feature pyramid. The experimental results showed that the mean average precision (mAP) and frame per second (FPS) of the proposed method reach 94.26% and 60.04%, which can lead most of the mainstream object detection networks in accuracy and speed and can be integrated into the mobile end to complete the task of high-precision and fast bridge detection.

### 17. AMFLW-YOLO: A Lightweight Network for Remote Sensing Image Detection

摘要: The scale of targets in remote sensing images varies greatly and is diverse. It has many small targets that are distributed densely and high complexity of image background. The number of network model parameters and the computation amount of the object detection algorithms based on deep learning is huge. It is difficult to apply them on a platform with fixed performance and limited computing resources. A lightweight remote sensing object detection model is proposed in this article, which called attention and multiscale feature fusion lightweight-YOLO (AMFLW-YOLO). The deep separable convolution, inverted residual, and linear bottleneck structure are employed to replace the standard convolution layer to reduce the model parameters in the backbone network of the model. The coordinate attention (CA) mechanism is introduced into the feature fusion network to capture the direction- and location-aware information across channels at the same time, which improves the accuracy of the network. The bidirectional feature pyramid network (BiFPN) structure is employed to strengthen feature extraction. The learnable weights are introduced to learn the importance of different input features. The multiscale feature fusion is applied to improve the detection effect. The experimental results show that the algorithm achieves satisfactory performance in terms of efficiency and accuracy and has advantages in detection accuracy and model lightweight.  
Web of Science 核心合集中的 "被引频次":  
13  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
13  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001102413500012  
Indexed Date:  
2023-12-11  
Institution Address:  
Tianjin Chengjian Univ, Sch Control & Mech Engn, Tianjin 300384, Peoples R China  
Institution Address:  
Harbin Inst Technol, Sch Instrument Sci & Engn, Harbin 150001, Peoples R China  
Institution:  
Institution:

### 18. Multiscale Feature Fusion and Threshold-Based Attentional YOLO for

摘要: Tailings ponds are a significant source of environmental pollution and present potential danger, making it a top priority to obtain accurate location information. Traditional object-based remote sensing techniques suffer from low efficiency and lack of automation. Furthermore, deep learning research on tailings ponds lacks practical models and datasets. To address these gaps, this letter proposes a novel approach-Multi-scale Feature Fusion and Threshold-based Attentional YOLO (MFTA-YOLO) for detecting tailings ponds in remote sensing images. Specifically, CSPDarknet53 is used as the feature extractor in MFTA-YOLO, and it combines the Multiscale Feature Fusion Module (MF) and a new attention mechanism Threshold-based Attention (TA). MF effectively integrates multiscale semantic features to detect varying sizes of objects. TA recalibrates input contributions based on the threshold gap, allowing better focus on regions and channels of interest. Additionally, a large-scale tailings ponds dataset, consisting of 3619 high-resolution (1.07 m) maps sourced from Google Earth, has been constructed. Extensive experiments on this dataset have demonstrated that MFTA-YOLO outperforms most existing methods for tailings pond detection in remote sensing images, achieving an F1-score of 0.799 and AP@.5 of 0.851.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001184884200008  
Indexed Date:  
2024-04-07  
Institution Address:  
China Univ Geosci, Sch Geog & Informat Engn, Wuhan 430074, Peoples R China  
Institution Address:  
Chinese Acad Sci, Innovat Acad Precis Measurement Sci & Technol, Wuhan 430071, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Resource & Environm Sci, Wuhan 430079, Peoples R China  
Institution:  
Institution:  
Institution:

### 19. MSFF-CDNet: A Multiscale Feature Fusion Change Detection Network for

摘要: Change detection (CD) is an important application of remote sensing (RS) technology, which discovers changes by observing bi-temporal RS images. The rise of deep learning provides new solutions for CD. However, due to the insufficient extraction and utilization of deep features from RS images, existing deep learning-based CD methods are difficult to fully integrate such deep features, resulting in unstable performance, especially low sensitivity to multiscale changes. In this letter, a multiscale feature fusion CD network (MSFF-CDNet) is proposed to enhance feature fusion, by integrating a mask-guided change fusion module (MGCF) to achieve the fusion of the consistency and difference of multiscale features. Also, a CD refinement module (CDRM) is implemented to assist the encoding-decoding structure to achieve CD at a finer scale. By training with a hybrid loss function, the MSFF-CDNet is able to learn transformation relationships of bi-temporal RS images and their change maps. Besides, using a deep supervised (DS) learning strategy further improves the fitting performance and robustness. The method is experimented on two open-source datasets (i.e., CDD and LEVIR-CD datasets). Compared to state-of-the-art (SOTA) CD methods, the proposed method outperforms all metrics and its intersection over union (IoU) reaches 92.39% and 85.89%, respectively.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
7  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001058816200012  
Indexed Date:  
2023-09-14  
Institution Address:  
China Univ Min & Technol, Sch Environm & Spatial Informat, Xuzhou 221116, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Remote Sensing & Informat Engn, Wuhan 430079, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Smart Cities Res Inst, Hong Kong, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Dept Land Surveying & Geoinformat, Hong Kong, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 20. A deep learning-based method for cervical transformation zone

摘要: BACKGROUND: Colposcopy is one of the common methods of cervical cancer screening. The type of cervical transformation zone is considered one of the important factors for grading colposcopic findings and choosing treatment. OBJECTIVE: This study aims to develop a deep learning-based method for automatic classification of cervical transformation zone from colposcopy images. METHODS: We proposed a multiscale feature fusion classification network to classify cervical transformation zone, which can extract features from images and fuse them at multiple scales. Cervical regions were first detected from original colposcopy images and then fed into our multiscale feature fusion classification network. RESULTS: The results on the test dataset showed that, compared with the state-of-the-art image classification models, the proposed classification network had the highest classification accuracy, reaching 88.49%, and the sensitivity to type 1, type 2 and type 3 were 90.12%, 85.95% and 89.45%, respectively, higher than the comparison methods. CONCLUSIONS: The proposed method can automatically classify cervical transformation zone in colposcopy images, and can be used as an auxiliary tool in cervical cancer screening.  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0928-7329  
eISSN:  
1878-7401  
入藏号:  
WOS:000949598700011  
PubMed ID:  
36093645  
Indexed Date:  
2023-03-31  
Institution Address:  
Tianjin Univ, Acad Med Engn & Translat Med, 92 Weijin Rd, Tianjin 300072, Peoples R China  
Institution Address:  
Tianjin Univ, Sch Precis Instrument & Optoelect Engn, Tianjin, Peoples R China  
Institution Address:  
Weifang Med Univ, Affiliated Hosp, Dept Obstet & Gynecol, Weifang, Shandong, Peoples R China  
Institution:  
Institution:  
Institution:

### 21. EFCOMFF-Net: A Multiscale Feature Fusion Architecture With Enhanced

摘要: Remote sensing images have the essential attribute of large-scale spatial variation and complex scene information, as well as the high similarity between various classes and the significant differences within the same class, which are easy to cause misclassification. To solve this problem, an efficient systematic architecture named multiscale feature fusion network with enhanced feature correlation (EFCOMFF-Net) is proposed to reduce the gap among multiscale features and fuse them to improve the representation ability of remote sensing images. First, to strengthen the correlation of multiscale features, a feature correlation enhancement module (FCEM) is specifically developed, which takes the features of different stages of the backbone network as input data to obtain multiscale features with enhanced correlation. Considering the differences between the shallow features and the deep features, the EFCOMFF-Net-v1 related to shallow features and EFCOMFF-Net-v2 related to deep features with different structures are proposed. Second, the designed two versions of the deep learning network focus on the global contour information and need to encode more accurate spatial information. A feature aggregation attention module (FAAM) is designed and embedded into the network to encode the deep features by applying the spatial information aggregation features. Finally, considering that the simple integration strategy cannot reduce the gap between the shallow multiscale features and the deep features, a feature refinement module (FRM) is presented to optimize the network. ResNet50, DenseNet121, and ResNet152 are selected to conduct a considerable number of experiments on four datasets, which show the superiority of our method compared to recent methods.  
在 BIOSIS Citation Index 中的被引频次:  
2  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
25  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
27  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000969953000016  
Indexed Date:  
2023-05-18  
Institution Address:  
Cent South Univ Forestry & Technol, Coll Comp & Informat Engn, Changsha 410004, Peoples R China  
Institution Address:  
Tokyo Inst Technol, Inst Innovat Res, Lab Future Interdisciplinary Res Sci & Technol, Tokyo 2268503, Japan  
Institution:  
Institution:

### 22. MFFSODNet: Multiscale Feature Fusion Small Object Detection Network for

摘要: Unmanned aerial vehicle (UAV) aerial image object detection is a valuable and challenging research field. Despite the breakthrough of deep learning-based object detection networks in natural scenes, UAV images often exhibit characteristics such as a high proportion of small objects, dense distribution, and significant variations in object scales, posing great challenges for accurate detection. To address these issues, we propose an innovative multiscale feature fusion small object detection network (MFFSODNet). First, concerning the high proportion of small objects in UAV images, an additional tiny object prediction head is introduced instead of the large object prediction head. This approach provides a good detection accuracy of small objects and significantly reduces the parameters. Second, to enhance the feature extraction capability of the network for fine-grained information from small objects, a multiscale feature extraction module (MSFEM) is designed, which could extract rich and valuable multiscale feature information through convolution operation of different scales on multiple branches. Third, to fuse the fine-grained information from shallow feature maps and the semantic information from deep feature maps, a new bidirectional dense feature pyramid network (BDFPN) is proposed. By expanding the feature pyramid network scale and introducing skip connections, BDFPN achieves efficient multiscale information fusion. Extensive experiments on the VisDrone and UAVDT benchmark datasets demonstrate that MFFSODNet outperforms the state-of-the-art object detection methods and further validate the effectiveness and generalization of MFFSODNet on photovoltaic array defect datasets (PVDs).  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
21  
被引频次合计:  
21  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001196700600020  
Indexed Date:  
2024-04-18  
Institution Address:  
Xijing Univ, Sch Elect Informat, Xian 710123, Peoples R China  
Institution Address:  
Xijing Univ, Xian Key Lab High Precis Ind Intelligent Vis Measu, Xian 710123, Peoples R China  
Institution Address:  
Shaanxi Jiurui Technol Co Ltd, Xian 710065, Shaanxi, Peoples R China  
Institution Address:  
Air Force Engn Univ, Air Traff Control & Ground Control Intercept Coll, Xian 710038, Peoples R China  
Institution Address:  
Xijing Univ, Sch Humanities & Educ, Xian 710123, Peoples R China  
Institution Address:  
Minzu Univ China, Sch Econ, Beijing 100081, Peoples R China  
Institution:  
Institution:  
Institution:  
Shaanxi Jiurui Technol Co Ltd  
Institution:  
Institution:  
Institution:

### 23. LSKF-YOLO: Large Selective Kernel Feature Fusion Network for Power Tower

摘要: With the rapid development of high-resolution satellite remote sensing observation technology, power tower detection based on satellite remote sensing images has become a key research focus for power intelligent inspection. However, the performance of power tower detection in satellite remote sensing images needs improvement due to complex backgrounds, and small and nonuniform target sizes. To address this, this article first constructs a multiscene high-resolution satellite remote sensing power tower dataset and then proposes the large selective kernel feature fusion (LSKF)-you only look once (YOLO) network for high-resolution satellite remote sensing images. This network primarily consists of a large spatial kernel selective attention fusion module and a multiscale feature alignment fusion (MFAF) structure. The large spatial selective kernel mechanism (LSKM) is improved by using the attentional feature fusion module, which provides richer feature information for accurately locating the position of the power tower. The MFAF structure effectively utilizes low-level semantic information, mitigates feature ambiguity in deeper network layers, and enables multiscale feature fusion of power towers within complex backgrounds. In addition, the introduction of minimum point distance-IoU (MPDIoU) enhances complete-IoU (CIoU), further improving the model's performance. The results demonstrate that the $F1$ score and mAP0.5 of the LSKF-YOLO network reach 0.764% and 77.47%, respectively. Compared with other deep learning-based satellite remote sensing power tower inspection methods, the LSKF-YOLO network significantly enhances detection accuracy and provides crucial technical support for intelligent inspection of power lines via satellite remote sensing.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
被引频次合计:  
6  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001225891900018  
Indexed Date:  
2024-06-08  
Institution Address:  
North China Elect Power Univ, Dept Elect & Commun Engn, Baoding 071003, Hebei, Peoples R China  
Institution Address:  
North China Elect Power Univ, Hebei Key Lab Power Internet Things Technol, Baoding 071003, Hebei, Peoples R China  
Institution:  
Institution:

### 24. Multiscale Feature Fusion for Gesture Recognition Using Commodity

摘要: Gestures are one of the most natural and intuitive approach for human-computer interaction. Compared with traditional camera-based or wearable sensors-based solutions, gesture recognition using the millimeter wave radar has attracted growing attention for its characteristics of contact-free, privacy-preserving and less environmentdependence. Although there have been many recent studies on hand gesture recognition, the existing hand gesture recognition methods still have recognition accuracy and generalization ability shortcomings in shortrange applications. In this paper, we present a hand gesture recognition method named multiscale feature fusion (MSFF) to accurately identify micro hand gestures. In MSFF, not only the overall action recognition of the palm but also the subtle movements of the fingers are taken into account. Specifically, we adopt hand gesture multiangle Doppler-time and gesture trajectory range-angle map multi-feature fusion to comprehensively extract hand gesture features and fuse high-level deep neural networks to make it pay more attention to subtle finger movements. We evaluate the proposed method using data collected from 10 users and our proposed solution achieves an average recognition accuracy of 99.7%. Extensive experiments on a public mmWave gesture dataset demonstrate the superior effectiveness of the proposed system.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001350270600041  
Indexed Date:  
2024-12-05  
Institution Address:  
Jinling Inst Technol, Coll Comp Engn, Nanjing 211169, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Coll Comp, Nanjing 210003, Peoples R China  
Institution:  
Institution:

### 25. SegR-Net: A deep learning framework with multi-scale feature fusion for

摘要: Retinal vessel segmentation is an important task in medical image analysis and has a variety of applications in the diagnosis and treatment of retinal diseases. In this paper, we propose SegR-Net, a deep learning framework for robust retinal vessel segmentation. SegR-Net utilizes a combination of feature extraction and embedding, deep feature magnification, feature precision and interference, and dense multiscale feature fusion to generate accurate segmentation masks. The model consists of an encoder module that extracts high-level features from the input images and a decoder module that reconstructs the segmentation masks by combining features from the encoder module. The encoder module consists of a feature extraction and embedding block that enhances by dense multiscale feature fusion, followed by a deep feature magnification block that magnifies the retinal vessels. To further improve the quality of the extracted features, we use a group of two convolutional layers after each DFM block. In the decoder module, we utilize a feature precision and interference block and a dense multiscale feature fusion block (DMFF) to combine features from the encoder module and reconstruct the segmentation mask. We also incorporate data augmentation and pre-processing techniques to improve the generalization of the trained model. Experimental results on three fundus image publicly available datasets (CHASE\_DB1, STARE, and DRIVE) demonstrate that SegR-Net outperforms state-of-the-art models in terms of accuracy, sensitivity, specificity, and F1 score. The proposed framework can provide more accurate and more efficient segmentation of retinal blood vessels in comparison to the state-of-the-art techniques, which is essential for clinical decision-making and diagnosis of various eye diseases.  
Web of Science 核心合集中的 "被引频次":  
36  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
8  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
37  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:001022601900001  
PubMed ID:  
37343468  
Indexed Date:  
2023-07-17  
Institution Address:  
Elect & Telecommun Res Inst, 176-11 Cheomdan Gwagi Ro, Gwangju 61012, South Korea  
Institution Address:  
Jeonbuk Natl Univ, Dept Elect & Informat Engn, Jeonju 54896, South Korea  
Institution Address:  
Bahria Univ, Dept Elect Engn, Islamabad, Pakistan  
Institution Address:  
Jeonbuk Natl Univ, Adv Elect & Informat Res Ctr, Jeonju 54896, South Korea  
Institution:  
Institution:  
Institution:  
Bahria Univ  
Institution:

### 26. A deep learning model for steel surface defect detection

摘要: Industrial defect detection is a hot topic in the field of computer vision. It is a challenging task due to complex features and many categories of industrial defects. In this paper, a deep learning model based on the multiscale feature extraction module is introduced for steel surface defect detection. The main focus on the feature extraction capability of the model and feature fusion capability to improve the accuracy of the model for steel surface defect detection. First, to improve the feature extraction ability of the model, a multiscale feature extraction (MSFE) module is introduced. The MSFE module can effectively extract multiscale features through three branches that have different convolution kernel sizes. Second, an efficient feature fusion (EFF) module is proposed to optimize feature fusion by adding features from the backbone network to the neck network. Third, this paper puts forward a new Bottleneck module by reducing the normalization layer and activation function in the original Bottleneck module. Finally, the backbone network is deepened to further enhance the feature extraction ability of the model. Extensive experiments are conducted on the public NEU-DET dataset. The experimental results validate the effectiveness of the designed modules and the proposed model. Compared with other state-of-the-art methods, the proposed model achieves optimal accuracy(73.08% mAP@0.5) while maintaining a small number of parameters.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
36  
被引频次合计:  
38  
ISSN:  
2199-4536  
eISSN:  
2198-6053  
入藏号:  
WOS:001046048100001  
Indexed Date:  
2023-08-24  
Institution Address:  
Qilu Univ Technol, Shandong Acad Sci, Key Lab Comp Power Network & Informat Secur, Shandong Comp Sci Ctr,Minist Educ, Jinan, Peoples R China  
Institution Address:  
Qilu Univ Technol, Shandong Acad Sci, Fac Comp Sci & Technol, Shandong Engn Res Ctr Big Dat Appl Technol, Jinan, Peoples R China  
Institution Address:  
Shandong Fundamental Res Ctr Comp Sci, Shandong Prov Key Lab Comp Networks, Jinan, Peoples R China  
Institution Address:  
Prince Sattam Bin Abdulaziz Univ, Coll Comp Engn & Sci, Dept Comp Sci, AlKharj 16278, Saudi Arabia  
Institution Address:  
South Valley Univ, Fac Comp & Informat, Dept Comp Sci, Qena, Egypt  
Institution Address:  
Queens Univ, Art & Sci Comp, Kingston, ON, Canada  
Institution:  
Institution:  
Institution:  
Shandong Fundamental Res Ctr Comp Sci  
Institution:  
Institution:  
Institution:

### 27. Adaptive Weighted Multi-Level Fusion of Multi-Scale Features: A New

摘要: Great achievements have been made in pedestrian detection through deep learning. For detectors based on deep learning, making better use of features has become the key to their detection effect. While current pedestrian detectors have made efforts in feature utilization to improve their detection performance, the feature utilization is still inadequate. To solve the problem of inadequate feature utilization, we proposed the Multi-Level Feature Fusion Module (MFFM) and its Multi-Scale Feature Fusion Unit (MFFU) sub-module, which connect feature maps of the same scale and different scales by using horizontal and vertical connections and shortcut structures. All of these connections are accompanied by weights that can be learned; thus, they can be used as adaptive multi-level and multi-scale feature fusion modules to fuse the best features. Then, we built a complete pedestrian detector, the Adaptive Feature Fusion Detector (AFFDet), which is an anchor-free one-stage pedestrian detector that can make full use of features for detection. As a result, compared with other methods, our method has better performance on the challenging Caltech Pedestrian Detection Benchmark (Caltech) and has quite competitive speed. It is the current state-of-the-art one-stage pedestrian detection method.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
4  
ISSN:  
1999-5903  
入藏号:  
WOS:000622599300001  
Indexed Date:  
2021-03-18  
Institution Address:  
Sichuan Univ, Coll Elect Engn, Chengdu 610065, Peoples R China  
Institution:

### 28. Cloud and Cloud Shadow Detection Tests Based on Multiscale Feature

摘要: Cloud detection based on high-resolution optical images combined with deep learning methodology cannot provide adequate and accurate information about the cloud, cloud shadows, or their edge details. The main reason for this problem is the insufficient fusion of semantic information in different scales of classification techniques. To address this problem, this study combines the Res.block (Residual block) module that can prevent network degradation, multiscale convolution module that can increase the receptive field of the network, and multiscale feature module that can extract and integrate information from different scales. In addition, this study proposes a detection algorithm based on the multiscale feature fusion network and deep learning. The experimental results showed that rich spatial and semantic information could be extracted by the algorithm. Cloud and cloud shadow masks with a higher level of accuracy can also be acquired. The accuracy of cloud and cloud shadow detection is 0.9351 and 0.8103, respectively. This study provides theoretical support and technical reserve for the application of deep learning techniques to operational cloud detection.  
摘要:  
基于深度学习的高分辨率光学影像云检测过程中,云和云阴影及其边缘细节丢失较为严重,主要原因在于不同尺度空间语义信息特征融合存在不足。针对该问题,本文构建一种基于深度学习的多尺度特征融合网络(Multi-scale Feature Fusion Network, MFFN)的云和云阴影检测方法,该算法结合防止网络退化的残差神经网络模块(Res.block)、扩大网络感受野的多尺度卷积模块(MCM)和提取并融合不同尺度信息的多尺度特征模块(MFM)。试验表明,本算法能提取丰富的空间信息与语义信息,可取得较为精细的云与云阴影掩模,具有较高检测精度,其中云检测准确率达0.9796,云阴影检测准确率达0.8307。同时,该工作可为深度学习技术应用于业务云检测提供理论支持及技术储备。  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-9895  
入藏号:  
CSCD:7096858  
Indexed Date:  
2022-02-08  
Institution Address:  
中国气象局国家卫星气象中心, 中国遥感卫星辐射测量和定标重点开放实验室, 北京 100081, 中国  
Institution Address:  
南京信息工程大学;;南京信息工程大学电子与信息工程学院, 江苏省大气环境与装备技术协同创新中心;;, 南京;;南京, ;; 210044;;210044  
Institution Address:  
中国科学院空间应用工程与技术中心, 北京 100094, 中国  
Institution Address:  
National Satellite Meteororologistic Center, Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites, China Meteorological Administration, Beijing 100081, China  
Institution Address:  
Nanjing University of Information Science and Technology;;School of Electronics and Information Engineering, Nanjing University of Information Science and Technology, Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology;;, Nanjing;;Nanjing, ;; 210044;;210044  
Institution Address:  
Technology and Engineering Center for Space Utilization, Chinese Academy of Sciences, Beijing 100094, China  
Institution:  
中国气象局国家卫星气象中心  
Institution:  
南京信息工程大学;;南京信息工程大学电子与信息工程学院  
Institution:  
中国科学院空间应用工程与技术中心  
Institution:  
National Satellite Meteororologistic Center  
Institution:  
Nanjing University of Information Science and Technology;;School of Electronics and Information Engineering, Nanjing University of Information Science and Technology  
Institution:  
Technology and Engineering Center for Space Utilization, Chinese Academy of Sciences

### 29. Real-time method for traffic sign detection and recognition based on

摘要: As a part of Intelligent Transportation System (ITS), the vehicle traffic sign detection and recognition system have been paid more attention by Intelligent transportation researchers, the traffic sign detection and recognition algorithm based on convolution neural network has great advantages in expansibility and robustness, but it still has great optimization space inaccuracy, computation and storage space. In this paper, we design a multiscale feature fusion algorithm for traffic sign detection and recognition. In order to improve the accuracy of the network, the gaussian distribution characteristics are used in the loss function. The training and analysis of two neural networks with different feature scales and YOLOv3-tiny were carried out on the Tsinghua-Tencent open traffic sign dataset. The experimental results show that the detection and recognition of the targets by networks with multiple feature scales have improved significantly, and the recall and accuracy are 95.32% and 93.13% respectively. Finally, the algorithm of traffic sign detection and recognition is verified on the NVIDIA Jetson Tx2 platform and delivers 28 fps outstanding performances.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
8  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
7  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
15  
ISSN:  
0954-4070  
eISSN:  
2041-2991  
入藏号:  
WOS:000643290400016  
Indexed Date:  
2021-05-28  
Institution Address:  
Shanghai Univ Engn Sci, 333 Longteng Rd, Shanghai 201620, Peoples R China  
Institution Address:  
Natl Inst Metrol, Beijing, Peoples R China  
Institution:  
Institution:

### 30. Multiscale fusion enhanced spiking neural network for invasive BCI

摘要: Brain-computer interfaces (BCIs) are an advanced fusion of neuroscience and artificial intelligence, requiring stable and long-term decoding of neural signals. Spiking Neural Networks (SNNs), with their neuronal dynamics and spike-based signal processing, are inherently well-suited for this task. This paper presents a novel approach utilizing a Multiscale Fusion enhanced Spiking Neural Network (MFSNN). The MFSNN emulates the parallel processing and multiscale feature fusion seen in human visual perception to enable real-time, efficient, and energy-conserving neural signal decoding. Initially, the MFSNN employs temporal convolutional networks and channel attention mechanisms to extract spatiotemporal features from raw data. It then enhances decoding performance by integrating these features through skip connections. Additionally, the MFSNN improves generalizability and robustness in cross-day signal decoding through mini-batch supervised generalization learning. In two benchmark invasive BCI paradigms, including the single-hand grasp-and-touch and center-and-out reach tasks, the MFSNN surpasses traditional artificial neural network methods, such as MLP and GRU, in both accuracy and computational efficiency. Moreover, the MFSNN's multiscale feature fusion framework is well-suited for the implementation on neuromorphic chips, offering an energy-efficient solution for online decoding of invasive BCI signals.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1662-453X  
入藏号:  
WOS:001439045100001  
PubMed ID:  
40061257  
Indexed Date:  
2025-03-13  
Institution Address:  
Chinese Acad Sci, Inst Automat, Beijing, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Artificial Intelligence, Beijing, Peoples R China  
Institution Address:  
Chinese Acad Sci, Ctr Excellence Brain Sci & Intelligence Technol, Shanghai, Peoples R China  
Institution:  
Institution:  
Institution:

### 31. MFFENet: Multiscale Feature Fusion and Enhancement Network For

摘要: Compared with traditional handcrafted features, deep learning has greatly improved the performance of scene parsing. However, it remains challenging under various environmental conditions caused by imaging limitations. Thermal imaging cameras have several advantages over cameras for the visible spectrum, such as operation in total darkness, robustness to shadow effects, insensitivity to illumination variations, and strong ability to penetrate smog and haze. These advantages of thermal imaging cameras make them ideal for the scene parsing of semantic objects in daytime and nighttime. In this paper, we propose a novel multiscale feature fusion and enhancement network (MFFENet) for accurate parsing of RGB-thermal urban road scenes even when the quality of the available RGB data is compromised. The proposed MFFENet consists of two encoders, a feature fusion layer, and a multi-label supervision layer. We concatenate the multi-scale features with the features that contain global semantic information. Furthermore, we explore the cross-modal fusion of RGB and thermal features at multiple stages, rather than fusing them once at the low or high stage. Then, we propose a spatial attention mechanism module that provides a higher weight to (focuses more on) the foreground area, allowing MFFENet to emphasize foreground objects. Finally, multi-label supervision is introduced to optimize parameters of the proposed MFFENet. Experimental results confirm that the proposed MFFENet outperforms similar high-performing methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
95  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
96  
ISSN:  
1520-9210  
eISSN:  
1941-0077  
入藏号:  
WOS:000793839600023  
Indexed Date:  
2022-05-29  
Institution Address:  
Zhejiang Univ Sci & Technol, Sch Informat & Elect Engn, Hangzhou 310023, Peoples R China  
Institution Address:  
Zhejiang Univ, Inst Informat & Commun Engn, Hangzhou 310027, Peoples R China  
Institution Address:  
Univ Washington, Dept Elect Engn, Seattle, WA 98105 USA  
Institution:  
Institution:  
Institution:

### 32. Building extraction based on multiple multiscale-feature fusion

摘要: A novel neural network named multiple multiscale-feature fusion attention network (MMFA-Net) was proposed for building segmentation from high-resolution remote sensing images aiming at the disadvantages that the fully convolutional networks for building extraction have the problems of over-segmentation and internal cavity.U-Net was used as the backbone combined with multiple-extract efficient channel attention (MECA) and multiscalefeature fusion attention (MFA) structure.The MECA module was designed to strengthen the effectiveness of the feature information through the weight ratio,which was in the skip connection.The transition allocation of attention to invalid features was avoided.The multiple feature extraction was adopted to reduce the loss of effective features.The MFA module was positioned at the bottom of the model.Different spatial features and spectral dimension features were obtained through the combination of parallel continuous medium or small-scale atrous convolution and channel attention.Then the problem of pixel loss of large buildings caused by atrous convolution was alleviated.The MMFA-Net integrating the MECA and the MFA modules can promote the integrity and accuracy of building extraction results.The proposed MMFA-Net was verified on WHU,Massachusetts,and owner-drawing building datasets.MMFA-Net showed better performance compared with the other five comparison methods.The F1-Score and IoU of MMFA-Net reached 93.33%,87.50% at WHU datasets,85.38%,74.49% at Massachusetts datasets,and 88.46%,79.31% at owner-drawing datasets,respectively.  
摘要:  
针对全卷积神经网络模型在进行建筑物提取时易产生过度分割以及内部空洞的问题,提出基于多重多尺度融合注意力网络(MMFA-Net)的高分辨率遥感影像建筑物提取方法.该方法以U-Net为主体架构,设计2个模块:多重高效通道注意力(MECA)和多尺度特征融合注意力(MFA).MECA设计在模型跳跃连接中,通过权重配比强化有效特征信息,避免注意力向无效特征的过渡分配;采用多重特征提取,减少有效特征的损失.MFA被嵌入模型底部,结合并行连续中小尺度空洞卷积与通道注意力,获得不同的空间特征与光谱维度特征,缓解空洞卷积造成的大型建筑物像素缺失问题.MMFA-Net通过融合MECA和MFA,提高了建筑物提取结果的完整度和精确率.将模型在WHU、Massachusetts和自绘建筑物数据集上进行验证,在定量评价方面优于其他5种对比方法,F1分数和IoU分别达到93.33%、87.50%;85.38%、74.49%和88.46%、79.31%.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1008-973X  
入藏号:  
CSCD:7330602  
Indexed Date:  
2023-02-25  
Institution Address:  
长江大学地球科学学院, 武汉, 湖北 430100, 中国  
Institution Address:  
中国科学院空天信息创新研究院;;海南省地球观测重点实验室, ;;海南省地球观测重点实验室, ;;三亚, 北京;;海南 100094;;572029, 中国  
Institution Address:  
长江大学地球科学学院;;湖南科技大学;;北京市测绘设计研究院, ;;测绘遥感信息工程湖南省重点实验室;;城市空间信息工程北京市重点实验室, 武汉;;湘潭;;, 湖北;;湖南;;北京 430100;;411201;;100045, 中国  
Institution Address:  
School of Geosciences,Yangtze University, Wuhan, Hubei 430100, China  
Institution Address:  
Aerospace Information Research Institute,Chinese Academy of Sciences;;Key Laboratory of Earth Observation of Hainan Province, ;;Key Laboratory of Earth Observation of Hainan Province, ;;Sanya, Beijing;; 100094;;572029  
Institution Address:  
School of Geosciences,Yangtze University;;Hunan University of Science and Technology;;Beijing Institute of Surveying and Mapping, ;;Hunan Provincial Key Laboratory of Geo-Information Engineering in Surveying,Mapping and Remote Sensing;;Beijing Key Laboratory of Urban Spatial Information Engineering, Wuhan;;Xiangtan;;, ;;;;Beijing 430100;;411201;;100045  
Institution:  
长江大学地球科学学院  
Institution:  
中国科学院空天信息创新研究院;;海南省地球观测重点实验室  
Institution:  
长江大学地球科学学院;;湖南科技大学;;北京市测绘设计研究院  
Institution:  
School of Geosciences,Yangtze University  
Institution:  
Aerospace Information Research Institute,Chinese Academy of Sciences;;Key Laboratory of Earth Observation of Hainan Province  
Institution:  
School of Geosciences,Yangtze University;;Hunan University of Science and Technology;;Beijing Institute of Surveying and Mapping

### 33. Multiscale Feature Fusion Method for Liver Cirrhosis Classification

摘要: Liver cirrhosis is one of the most common liver diseases in the world, posing a threat to people's daily lives. In advanced stages, cirrhosis can lead to severe symptoms and complications, making early detection and treatment crucial. This study aims to address this critical healthcare challenge by improving the accuracy of liver cirrhosis classification using ultrasound imaging, thereby assisting medical professionals in early diagnosis and intervention. This article proposes a new multiscale feature fusion network model (MSFNet), which uses the feature extraction module to capture multiscale features from ultrasound images. This approach enables the neural network to utilize richer information to accurately classify the stage of cirrhosis. In addition, a new loss function is proposed to solve the class imbalance problem in medical datasets, which makes the model pay more attention to the samples that are difficult to classify and improves the performance of the model. The effectiveness of the proposed MSFNet was evaluated using ultrasound images from 61 subjects. Experimental results demonstrate that our method achieves high classification accuracy, with 98.08% on convex array datasets and 97.60% on linear array datasets. Our proposed method can classify early, middle, and late cirrhosis very accurately. It provides valuable insights for the clinical treatment of liver cirrhosis and may be helpful for the rehabilitation of patients.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0899-9457  
eISSN:  
1098-1098  
入藏号:  
WOS:001268742900001  
Indexed Date:  
2024-07-23  
Institution Address:  
Hubei Univ Technol, Sch Elect & Elect Engn, Wuhan, Peoples R China  
Institution Address:  
Tongji Hosp, Tongji Med Coll HUST, Wuhan, Peoples R China  
Institution:  
Institution:

### 34. LiM-Net: Lightweight multi-level multiscale network with deep residual

摘要: Automatic liver segmentation gained significant attention in the medical realm to deal with liver anomalies. Furthermore, due to advancements in medical imaging, data volume is increasing day-to-day, which seeks the demand for automatic liver segmentation techniques to evade the labour-intensive process of liver delineation currently followed by medical experts. The proposed method is based on the deep learning approach. We exploited the multi-level multiscale feature extraction and fusion concept to uplift the liver segmentation outcome. The computationally efficient pre-activated multiscale Res2Net backbone architecture with channelwise attention (PARCA) block plugged into the Unet++ architecture to extract the multiscale fine-grained features with refinement and dense skip connections used for the multiscale feature fusion from the various stages of the network. As a result, the fine-grained multiscale features and multiscale feature fusion from diverse stages provide rich contextual feature representation that enhances decoder competence. Further, we optimized the network using a custom loss function that handles the class imbalance and focuses on the complicated samples from the dataset. The efficacy of the proposed lightweight model was tested experimentally using the publicly available 3DIRCADb, CHAOS and LiTS CT datasets.The proposed model achieved the DSC of 97.3%, 95.1%, and 96.3% on the 3DIRCADb, CHAOS, and LiTS datasets. Nevertheless, the proposed network is lightweight and has 7.5 million parameters which are less than the classical Unet and Unet++ architecture. Thus, the proposed heuristics uplift the liver segmentation outcome and significantly reduce the parameters and the model's computational complexity.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
22  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
22  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:000878640700004  
Indexed Date:  
2022-11-13  
Institution Address:  
Shri Guru Gobind Singhji Inst Engn & Technol, Dept Elect & Telecommun Engn, Nanded, Maharashtra, India  
Institution:

### 35. OAMSFNet: Orientation-Aware and Multi-Scale Feature Fusion Network for

摘要: Remote sensing images (RSI) of urban regions often exhibit shadows cast by buildings and other objects, which may result in imprecise analysis and interpretation. Therefore, shadow detection plays a significant role in RSI scene understanding. Current approaches have paid little attention to the presence of pseudo shadows and their confounding effects on detection results. We tackle these questions in the spatial context of orientation-awareness, effectively capturing the intricate relationships between shadows and ground objects. Here, we introduce a novel deep learning network named the Orientation-Aware and Multi-Scale Feature Fusion Network (OAMSFNet) due to noise reduction considerations. The proposed OAMSFNet comprises Shadow Aware Feature Encoder (SFE), Orientation-Aware Context Module (OCM), and Multi-Scale Feature Pyramid (MFP). Furthermore, a multi-scale feature fusion algorithm has been devised to enhance the detection and segmentation capabilities of the model in shadow regions. Finally, a comparative study was conducted on the Aerial Imagery dataset for Shadow Detection (AISD) in both quantitative and qualitative aspects. The experimental results show that our method outperforms state-of-the-art methods while maintaining a lightweight model design, indicating that our approach exhibits remarkable accuracy and stability in excellent agreement with predictions. OAMSFNet achieved an average F-score of 89.85%, surpassing the straightforward semantic segmentation model SegNet by 9.94%, and demonstrated remarkable efficiency enhancements, with FLOPs reduced to as low as 11.48 G, resulting in efficiency improvements of 66.7%-96.5% compared to other shadow detection models.  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:001058786100001  
Indexed Date:  
2023-09-14  
Institution Address:  
Chongqing Univ Posts & Telecommun, Coll Comp Sci & Technol, Chongqing, Peoples R China  
Institution Address:  
Pearl River Water Resources Commiss, Pearl River Water Resources Res Inst, Guangzhou, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Coll Comp Sci & Technol, Chongqing 400065, Peoples R China  
Institution:  
Institution:  
Pearl River Water Resources Commiss  
Institution:

### 36. Tunnel Crack Detection With Linear Seam Based on Mixed Attention and

摘要: Crack detection techniques have been rapidly developed in recent years due to the rise of deep learning. However, existing methods struggle to produce accurate crack segmentation results because cracks and linear seams on the tunnel lining surface have significant similarities in terms of intensity value and texture features. At the same time, due to the scarcity of data, the existing tunnel lining surface crack detection methods still use multistep traditional image processing methods for detection, which is inefficient. In this article, we collect and label a dataset of 200 tunnel lining surface crack images named Tunnel200. For the first time, a deep-learning-based method is used to detect cracks in the tunnel lining surface. To deal with the characteristics of crack and linear seam, which mostly present long strip or curved shapes, we propose a mixed attention (MA) module by efficient embedding channel and positional information. Unlike common spatial attention that aggregates information throughout space, MA aggregates feature directly along with two directions, height, and width, in the spatial dimension. In this way, the long-range dependence of the crack features can be effectively captured. The proposed MA is simple to incorporate into the network. Meanwhile, we embed it in the traditional U-shape network while using an efficient multiscale feature fusion technique to build the tunnel crack detection network (TCDNet). TCDNet outperforms other crack detection and semantic segmentation methods on the Tunnel200 dataset. In addition, we evaluate our method on two publicly available crack datasets, Crack500 and DeepCrack, and our method gets superior performance.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
37  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
38  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000838416300004  
Indexed Date:  
2022-08-21  
Institution Address:  
Chongqing Univ Posts & Telecommun, Sch Comp Sci & Technol, Chongqing 400065, Peoples R China  
Institution:

### 37. A multispectral feature fusion network for robust pedestrian detection

摘要: The multispectral information, including both visible information and infrared information, can describe the detection target in a comprehensive manner. The deep learning (DL)-based detectors that fuse the multispectral features can detect pedestrians robustly in various environments. Therefore, this paper puts forward a robust multispectral feature fusion network (MSFFN) for pedestrian detection, which fully integrates the features extracted from visible light and infrared channels. Specifically, multiscale semantic features were extracted by two core modules, namely, multiscale feature extraction of visible images (MFEV) and multiscale feature extraction of infrared images (MFEI), and fused by the improved YOLOv3 network for pedestrian recognition. Through experiments on the KAIST dataset, it is proved that the MSFFN model can detect pedestrians more accurately than both MFEV and MFEI over daytime and nighttime images on multiple scales. The experimental results on the KAIST multispectral data set in the last section showed that our proposed MFMFN model was superior to a number of state-of-the-art multispectral pedestrian detectors methods in accuracy and speed. The model was also found to strike a good balance between accuracy and speed, and perform excellently on small input images. The research results shed important new light on the design of self-driving vehicles. (C) 2020 The Authors. Published by Elsevier B.V. on behalf of Faculty of Engineering, Alexandria University.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
31  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
33  
ISSN:  
1110-0168  
eISSN:  
2090-2670  
入藏号:  
WOS:000604991100007  
Indexed Date:  
2021-01-22  
Institution Address:  
Xian Technol Univ, Sch Elect & Informat Engn, Xian 710021, Peoples R China  
Institution:

### 38. Multi-scale feature fusion based DOA and range estimation for near-field

摘要: In this paper, a multi-scale deep neural network structure is proposed to estimate the direction-of-arrivals (DOAs) and range parameters for near-field source localization. Initially, the covariance matrix is used as the network input for feature extraction through different convolutional operations, obtaining multi-scale information on DOA and range. Subsequently, the quantum genetic algorithm is employed to perform a weighted fusion of multi-scale features, resulting in the output near-field source localization parameters. Finally, we analyzed the relevant parameters of multi-scale feature extraction and feature-weighted fusion. Numerical simulation results demonstrate that the proposed method exhibits superior performance in near-field source parameter estimation compared to traditional methods.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001365843600003  
Indexed Date:  
2024-12-05  
Institution Address:  
Harbin Univ Sci & Technol, Higher Educ Key Lab Measuring & Control Technol &, Harbin 150080, Heilongjiang, Peoples R China  
Institution Address:  
Harbin Univ Sci & Technol, Sch Automat, Harbin 150080, Peoples R China  
Institution:  
Institution:

### 39. Visual Image Decoding of Brain Activities Using a Dual Attention

摘要: Reconstructing visual stimulus from human brain activity measured with functional magnetic resonance imaging (fMRI) is a challenging decoding task for revealing the visual system. Recent deep learning approaches commonly neglect the relationship between hierarchical image features and different regions of the visual cortex, and fail to use global and local image features in reconstructing visual stimulus. To address these issues, in this article, a novel neural decoding framework is proposed by using a dual attention (DA) hierarchical latent generative network with multiscale feature fusion (DA-HLGN-MSFF) method. Specifically, the fMRI data are first encoded to hierarchical features of our image encoder network, which employs a multikernel convolution block to extract the multiscale spatial information of images. In order to reconstruct the perceived images and further improve the performance of our generator network, a DA block based on the channel-spatial attention mechanism is then proposed to exploit the interchannel relationships and spatial long-range dependencies of features. Moreover, a multiscale feature fusion block is finally adopted to aggregate the global and local information of features at different scales and synthesize the final reconstructed images in the generator network. Competitive experimental results on two public fMRI data sets demonstrate that our method is able to achieve promising reconstructing performance compared with the state-of-the-art methods. The codes of our proposed DA-HLGN-MSFF method will be open access on https://github.com/ljbuaa/HLDAGN.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
2379-8920  
eISSN:  
2379-8939  
入藏号:  
WOS:001005746000039  
Indexed Date:  
2023-07-18  
Institution Address:  
Beihang Univ, Sch Automat Sci & Elect Engn, Beijing 100191, Peoples R China  
Institution Address:  
Beihang Univ, Sch Engn Med, Beijing 100191, Peoples R China  
Institution Address:  
Beijing Aerosp Automat Control Inst, Natl Key Lab Sci & Technol Aerosp Intelligence Con, Beijing 100070, Peoples R China  
Institution:  
Institution:  
Institution:  
Beijing Aerosp Automat Control Inst

### 40. Bi-Branch Multiscale Feature Joint Network for ORSI Salient Object

摘要: Salient object detection (SOD) of optical remote sensing images (ORSIs) has been a crucial part of the remote sensing field. In recent years, with the development of deep learning, many salient detection models for ORSIs have emerged. However, current study is limited to sunny weather conditions, and there is a lack of research on SOD in adverse weather conditions. Traditional models lack robustness and tend to miss detection in adverse weather conditions. To address this challenge, this article proposes a bi-branch multiscale feature joint network (BMFJNet) that achieves SOD in adverse weather conditions through a bi-branch linear joint structure. First, we obtain clean ORSIs through the dark channel prior and feed the clean images and the hazy images by two linear branches to the backbone for feature extraction, respectively. Second, the obtained effective features are input to the detection module for salient analysis. The detection module consists of three key components, where the multiscale feature aggregation module (MFAM) achieves salient feature enhancement in both dimensional directions through an attention mechanism, the adjacent pooling guidance module (APGM) guides the contextual information of adjacent layers through multiple pooling layers, and the feature fusion module aggregates global information from different components. In addition, we introduce a self-supervised robust restoration loss that enables our network to cope with different levels of adverse weather. Extensive experiments on synthetic datasets demonstrate the superiority of our proposed model over other state-of-the-art models on various metrics.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001346188500005  
Indexed Date:  
2024-11-26  
Institution Address:  
Southwest Univ, Coll Artificial Intelligence, Chongqing 400715, Peoples R China  
Institution:

### 41. Flooded Infrastructure Change Detection in Deeply Supervised Networks

摘要: Flood disasters are frequent, sudden, and have significant chain effects, seriously damaging infrastructure. Remote sensing images provide a means for timely flood emergency monitoring. When floods occur, emergency management agencies need to respond quickly and assess the damage. However, manual evaluation takes a significant amount of time; in current, commercial applications, the post-disaster flood vector range is used to directly overlay land cover data. On the one hand, land cover data are not updated in time, resulting in the misjudgment of disaster losses; on the other hand, since buildings block floods, the above methods cannot detect flooded buildings. Automated change-detection methods can effectively alleviate the above problems. However, the ability of change-detection structures and deep learning models for flooding to characterize flooded buildings and roads is unclear. This study specifically evaluated the performance of different change-detection structures and different deep learning models for the change detection of flooded buildings and roads in very-high-resolution remote sensing images. At the same time, a plug-and-play, multi-attention-constrained, deeply supervised high-dimensional and low-dimensional multi-scale feature fusion (MSFF) module is proposed. The MSFF module was extended to different deep learning models. Experimental results showed that the embedded MSFF performs better than the baseline model, demonstrating that MSFF can be used as a general multi-scale feature fusion component. After FloodedCDNet introduced MSFF, the detection accuracy of flooded buildings and roads changed after the data augmentation reached a maximum of 69.1% MIoU. This demonstrates its effectiveness and robustness in identifying change regions and categories from very-high-resolution remote sensing images.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001366089100001  
Indexed Date:  
2024-12-05  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Minist Emergency Management, Key Lab Emergency Satellite Engn & Applicat, Beijing 100124, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100094, Peoples R China  
Institution Address:  
Natl Disaster Reduct Ctr China, 6 Guangbaidong Rd, Beijing 100124, Peoples R China  
Institution:  
Institution:  
Minist Emergency Management  
Institution:  
Institution:  
Natl Disaster Reduct Ctr China

### 42. Multisource Multimodal Feature Fusion for Small Leak Detection in Gas

摘要: Improving the ability to detect small leaks to prevent more severe accidents plays an extremely important role in the safe operation of pipelines. To tackle the issue of low diagnostic accuracy associated with single sensors for detecting small leaks, a multisource multimodal feature fusion method for gas pipeline leak detection was proposed. First, the collected data from multiple sensors were transformed into 2-D time-frequency images for input into the feature extraction network. Then, the dual-information fusion (DIF) module was introduced, incorporating the attention mechanism and multiscale feature fusion to enhance the model's feature expression capability and fully interact with the multimodal features. Second, the channel split multiscale convolution (CSMC) module was designed to accommodate the diversity of input data and improve the model's generalization capability. The DIF and CSMC modules were cascaded and fused to produce the classification results through the fully connected layer. Finally, the effectiveness of the proposed method was assessed using pipeline leak data collected in the laboratory. The experimental results demonstrate that the proposed multimodal deep learning model can effectively identify the small leak state in pipelines, exhibiting superior diagnostic performance when compared to the current mainstream image classification models.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
9  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001166991700098  
Indexed Date:  
2024-04-04  
Institution Address:  
Northeast Petr Univ, SANYA Offshore Oil & Gas Res Inst, Sanya 572024, Peoples R China  
Institution Address:  
Coll Elect & Informat Engn, Daqing 163318, Peoples R China  
Institution Address:  
Qilu Inst Technol, Sch Comp & Informat Engn, Jinan 250200, Peoples R China  
Institution Address:  
China Petr Mat Co Ltd, Beijing 100032, Peoples R China  
Institution:  
Institution:  
Coll Elect & Informat Engn  
Institution:  
Qilu Inst Technol  
Institution:  
China Petr Mat Co Ltd

### 43. Damper defect detection for transmission line based on cognitive

摘要: Defect detection based on deep learning has been applied in many fields and has made great progress in recent years. However, the problems of high false detection rate in complex scenes and weak multiscale detection capability have not been solved well. To solve the above problems, a damper defect detection network (DDDNet) is proposed, which includes a cognitive preprocessing network (CogniPrepNet) and a heterogeneous feature pyramid network (HeteroFPN). CogniPrepNet enhances targets and mitigates the impact of complex backgrounds through the ghost-shuffle module. HeteroFPN obtains feature maps with a wider range of scales through a many-to-one mapping relationship to achieve high-quality multiscale feature fusion. Moreover, a fast proportional intersection over union is designed to improve the sensitivity of geometric factors during the bounding box regression. The experimental results on damper-DET dataset show that the DDDNet can achieve 98.25% mAP and 34 FPS detection speed in complex background and multiscale target scene. (c) 2023 SPIE and IS&T  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:001172000900008  
Indexed Date:  
2024-03-17  
Institution Address:  
South China Univ Technol, Sch Automat Sci & Engn, Guangzhou, Peoples R China  
Institution:

### 44. WFF-Net: Trainable weight feature fusion convolutional neural networks

摘要: Deep learning-based surface defect segmentation (SDS) technique is widely used in the field of surface defect detection (SDD) for its high accuracy and robustness. However, the deep learning-based surface defect segmentation method suffers from the interference from the semantic difference between the high-dimensional features and the low-dimensional features during the process of multi-scale feature fusion, which will bring additional noise to the network and thus affect the detection accuracy. For this drawback, this paper investigates a new weight-based feature fusion method, which aims to reduce the semantic differences and information redundancy after multiscale feature fusion in the process of coupling high-dimensional semantic features with low-dimensional semantic features. First, the WFF feature fusion method is proposed in the multi-scale feature fusion stage, which uses a learnable Gate module to assign weight coefficients to neighboring features, and uses an attention mechanism to fuse weighted neighboring features, so that redundant information can be reduced both before neighboring features are coupled. It can also reduce the semantic differences between multi-highdimensional features and low-dimensional features after fusion. Second, a dual decoding module is constructed to reduce the feature loss in the decoding stage, and a structural loss function is designed to optimize the network for the multi-scale output in the dual decoder. The proposed WFF-Net has been conducted on three datasets, and it shows that the proposed WFF-Net outperforms several existing DL methods in mean intersection of union (NEU-SEG: 85.70%, DAGM 2007: 86.12%, MT defects: 82.72%) and F1-measure (NEU-SEG: 94.11 %, DAGM 2007: 96.32 %, MT defects: 93.90 %).  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1474-0346  
eISSN:  
1873-5320  
入藏号:  
WOS:001394431700001  
Indexed Date:  
2025-01-17  
Institution Address:  
China Univ Geosci, Sch Mech Engn & Elect Informat, 388 LuMo Rd, Wuhan 430074, Peoples R China  
Institution Address:  
China Univ Geosci, Shenzhen Res Inst, Shenzhen 518057, Peoples R China  
Institution Address:  
Huazhong Univ Sci & Technol, State Key Lab Digital Mfg Equipment & Technol, 1037 Luoyu Rd, Wuhan 430074, Peoples R China  
Institution Address:  
Chongqing Univ, State Key Lab Mech Transmiss, Chongqing 400023, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 45. Abnormal behavior recognition based on feature fusion C3D network

摘要: The real-time detection and recognition ability of human action recognition in a video surveillance system is a key problem in an intelligent surveillance system. Because the behavior recognition for video surveillance systems is affected by the complexity of the scene, the classification performance of the behavior recognition models is not satisfactory. To increase the processing efficiency of the network and solve the problem of low classification accuracy of human action recognition, we designed a deep learning model based on three-dimensional (3D) convolutional network multiscale feature fusion to reduce the impact of constant appearance changes, background clutter, and pedestrian occlusion. The model alternately uses 3D convolution and 3D pooling operations to extract temporal and spatial features between consecutive frames after data preprocessing, and then uses a feature pyramid structure to select three sets of feature layers with different scales. The model performs deconvolution operations in a bottom to up order and fuses with the features of the previous layer, then downsampling and high-level feature layer fusion are performed sequentially from top to bottom. Using the newly generated the highest-level feature layer to realize abnormal behavior recognition. The C3D network algorithm based on feature fusion proposed in this paper is compared with the three most advanced methods of C3D, R3D, and R ( 2 + 1 ) D on the pedestrian abnormal action recognition (PAAR) dataset and the same parameters, and the accuracy is significantly improved.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:000981645700012  
Indexed Date:  
2023-06-13  
Institution Address:  
Zhengzhou Univ Light Ind, Sch Comp & Commun Engn, Zhengzhou, Peoples R China  
Institution Address:  
Zhengzhou Univ Light Ind, Henan Key Lab Food Safety Data Intelligence, Zhengzhou, Peoples R China  
Institution Address:  
Simshine Intelligent Technol Co Ltd, Ningbo, Peoples R China  
Institution Address:  
Henan Hongbo Measurement & Control Technol Co Ltd, Zhengzhou, Peoples R China  
Institution:  
Institution:  
Institution:  
Simshine Intelligent Technol Co Ltd  
Institution:  
Henan Hongbo Measurement & Control Technol Co Ltd

### 46. Sparse-View Photoacoustic Reconstruction Method for Diabetic Retinopathy

摘要: Diabetic retinopathy is one of the most prevalent microvascular complications of diabetes mellitus, and photoacoustic imaging is an effective method for imaging diabetic retinal vessels. Photoacoustic imaging is an emerging noninvasive imaging method based on the photoacoustic effect, which offers advantages of contrast, resolution, and depth imaging. Appropriate photoacoustic reconstruction methods are essential for obtaining high-quality photoacoustic images. In this study, a multi-input self-attention multiscale feature fusion network (SAMF-Net) is proposed for photoacoustic reconstruction. The algorithm accepts two inputs, namely the original photoacoustic signal and the traditional reconstructed image. Furthermore, a global feature extraction module based on the self-attention mechanism is employed to focus on the global information. The results demonstrate that the proposed method exhibits superior reconstruction capability under different sparse detection views. The method has instructive value for photoacoustic image reconstruction and has the potential for further application in the diagnosis of diabetic retinopathy.  
To address the issue of image quality degradation caused by under-sampling in photoacoustic imaging systems, a deep learning reconstruction algorithm called self-attentive multiscale feature fusion network is proposed. The algorithm accepts two inputs, namely the original photoacoustic signal and the traditional reconstructed image. Furthermore, a global feature extraction module based on the self-attention mechanism is employed. The algorithm significantly enhances the performance of photoacoustic reconstruction under sparse viewpoints.image  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1864-063X  
eISSN:  
1864-0648  
入藏号:  
WOS:001328137200001  
PubMed ID:  
39379333  
Indexed Date:  
2024-10-12  
Institution Address:  
Shandong Univ, Ctr Opt Res & Engn, Qingdao, Peoples R China  
Institution Address:  
China Univ Min & Technol, Sch Informat & Control Engn, Xuzhou, Jiangsu, Peoples R China  
Institution:  
Institution:

### 47. GDRS-YOLO: More Efficient Multiscale Features Fusion Object Detector for

摘要: The topic of object detection (OD) in remote sensing (RS) has received a lot of attention due to the rapid growth of deep learning. However, RS images typically have the following characteristics: significant variations in object scales, tight arrangement of small objects, and indistinguishable feature boundaries between objects and backgrounds. These challenges lead to defects, such as insufficient feature extraction and information loss of the existing methods. To address the above issues, based on the YOLOv7 architecture, we present a novel OD method named GDRS-you only look once (YOLO). Our primary contributions include: first, an enhanced feature extraction network based on deformable convolution is proposed to improve the network's ability to model geometric transformations. Second, we abandoned the traditional feature pyramid architecture and construct a multiscale feature aggregation network based on the gather-and-distribute mechanism, which makes effective use of the feature obtained from the backbone and reduces the loss of information in the transmission process. Finally, the normalized Wasserstein distance (NWD) is introduced for hybrid loss training, which alleviates the sensitivity of the IoU-based metric to the location deviation of tiny objects. We demonstrate the effectiveness of GDRS-YOLO on the publicly available datasets NWPU VHR-10 and VisDrone datasets. Compared to the original YOLOv7, the proposed method improves the mean average precision (mAP) by 1.9% and 5.5%, respectively. These results emphasize the superior performance of the proposed model, which provides an efficient multiscale feature fusion solution for RS applications.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001271937700014  
Indexed Date:  
2024-07-26  
Institution Address:  
Nanjing Univ Sci & Technol, Sch Automat, Nanjing 210094, Peoples R China  
Institution Address:  
Shanghai Electromech Engn Inst, Shanghai 201109, Peoples R China  
Institution:  
Institution:  
Shanghai Electromech Engn Inst

### 48. Retinal optical coherence tomography image classification based on

摘要: The retinal optical coherence tomography(OCT) image classification method based on deep learning has problems such as low ability of network feature extraction and difficult classification of small target lesions. Therefore, this paper proposes a dual branch multiscale feature fusion network. The gating attention mechanism is added to the vgg16 network, and the deep features are transmitted to the shallow features as gating signals. The redundant features are removed more fine-grained abstract information is obtained. Simultaneously, an atrous spatial pyramid pooling(ASPP) module is introduced to increase the receptive field and capture the global context information in various proportions without reducing the feature map resolution. The ASPP module increases the classification accuracy of small target lesions. The experimental results show that the proposed method has achieved good results in the retinal OCT image classification task, and the classification accuracy has reached 97.9%.  
摘要:  
目前基于深度学习的视网膜OCT图像分类方法存在网络特征提取能力低、小目标病变分类困难等问题。为此本文提出了一种双分支多尺度特征融合网络,通过加入门控注意力机制,利用深层特征作为选通信号传递给浅层特征,在消除冗余特征的同时,获得更细尺度的抽象信息。同时加入空洞空间金字塔模块,实现在不降低特征图分辨率的同时增大感受野,按不同比例有效捕获全局上下文信息,提高了小目标病变分类精度。实验结果表明,本文提出的方法在视网膜OCT图像分类任务中取得了较好效果,分类准确率达97.9%。  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
1673-4785  
入藏号:  
CSCD:7237708  
Indexed Date:  
2022-09-09  
Institution Address:  
哈尔滨工程大学信息与通信工程学院, 哈尔滨, 黑龙江 150001, 中国  
Institution Address:  
中央民族大学信息工程学院, 北京 100081, 中国  
Institution Address:  
College of Information and Communication Engineering, Harbin Engineering University, Harbin, Heilongjiang 150001, China  
Institution Address:  
School of Information Engineering, Minzu University of China, Beijing 100081, China  
Institution:  
哈尔滨工程大学信息与通信工程学院  
Institution:  
中央民族大学信息工程学院  
Institution:  
College of Information and Communication Engineering, Harbin Engineering University  
Institution:  
School of Information Engineering, Minzu University of China

### 49. LRSNet: a high-efficiency lightweight model for object detection in

摘要: . Unmanned aerial vehicles (UAVs) exhibit the ability to flexibly conduct aerial remote-sensing imaging. By employing deep learning object-detection algorithms, they efficiently perceive objects, finding widespread application in various practical engineering tasks. Consequently, UAV-based remote sensing object detection technology holds considerable research value. However, the background of UAV remote sensing images is often complex, with varying shooting angles and heights leading to difficulties in unifying target scales and features. Moreover, there is the challenge of numerous densely distributed small targets. In addition, UAVs face significant limitations in terms of hardware resources. Against this background, we propose a lightweight remote sensing object detection network (LRSNet) model based on YOLOv5s. In the backbone of LRSNet, the lightweight network MobileNetV3 is used to substantially reduce the model's computational complexity and parameter count. In the model's neck, a multiscale feature pyramid network named CM-FPN is introduced to enhance the detection capability of small objects. CM-FPN comprises two key components: C3EGhost, based on GhostNet and efficient channel attention modules, and the multiscale feature fusion channel attention mechanism (MFFC). C3EGhost, serving as CM-FPN's primary feature extraction module, possesses lower computational complexity and fewer parameters, as well as effectively reducing background interference. MFFC, as the feature fusion node of CM-FPN, can adaptively weight the fusion of shallow and deep features, acquiring more effective details and semantic information for object detection. LRSNet, evaluated on the NWPU VHR-10, DOTA V1.0, and VisDrone-2019 datasets, achieved mean average precision of 94.0%, 71.9%, and 35.6%, with Giga floating-point operations per second and Param (M) measuring only 5.8 and 4.1, respectively. This outcome affirms the efficiency of LRSNet in UAV-based remote-sensing object detection tasks.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
1931-3195  
入藏号:  
WOS:001271750200040  
Indexed Date:  
2024-08-04  
Institution Address:  
Beijing Informat Sci & Technol Univ, Sch Commun, Beijing, Peoples R China  
Institution Address:  
Minist Educ Optoelect Measurement Technol & Instr, Key Lab, Beijing, Peoples R China  
Institution:  
Institution:  
Minist Educ Optoelect Measurement Technol & Instr

### 50. Modulation recognition network of multi-scale analysis with deep

摘要: To improve the accuracy of modulated signal recognition in variable environments and reduce the impact of factors such as lack of prior knowledge on recognition results, researchers have gradually adopted deep learning techniques to replace traditional modulated signal processing techniques. To address the problem of low recognition accuracy of the modulated signal at low signal-to-noise ratios, we have designed a novel modulation recognition network of multi-scale analysis with deep threshold noise elimination to recognize the actually collected modulated signals under a symmetric cross-entropy function of label smoothing. The network consists of a denoising encoder with deep adaptive threshold learning and a decoder with multi-scale feature fusion. The two modules are skip-connected to work together to improve the robustness of the overall network. Experimental results show that this method has better recognition accuracy at low signal-to-noise ratios than previous methods. The network demonstrates a flexible self-learning capability for different noise thresholds and the effectiveness of the designed feature fusion module in multi-scale feature acquisition for various modulation types.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
2095-9184  
eISSN:  
2095-9230  
入藏号:  
WOS:001000508800009  
Indexed Date:  
2023-06-16  
Institution Address:  
Harbin Engn Univ, Coll Informat & Commun Engn, Harbin 150001, Peoples R China  
Institution Address:  
Harbin Engn Univ, Key Lab Adv Marine Commun & Informat Technol, Minist Ind & Informat Technol, Harbin 150001, Peoples R China  
Institution Address:  
China Coal Technol Engn Grp Chongqing Res Inst, Chongqing 400037, Peoples R China  
Institution Address:  
State Key Lab Methane Disaster Monitoring & Emerge, Chongqing 400039, Peoples R China  
Institution:  
Institution:  
Institution:  
China Coal Technol Engn Grp Chongqing Res Inst  
Institution:  
State Key Lab Methane Disaster Monitoring & Emerge

### 51. Cross-SRN: Structure-Preserving Super-Resolution Network With Cross

摘要: It is challenging to restore low-resolution (LR) images to super-resolution (SR) images with correct and clear details. Existing deep learning works almost neglect the inherent structural information of images, which acts as an important role for visual perception of SR results. In this paper, we design a hierarchical feature exploitation network to probe and preserve structural information in a multi-scale feature fusion manner. First, we propose a cross convolution upon traditional edge detectors to localize and represent edge features. Then, cross convolution blocks (CCBs) are designed with feature normalization and channel attention to consider the inherent correlations of features. Finally, we leverage multi-scale feature fusion group (MFFG) to embed the cross convolution blocks and develop the relations of structural features in different scales hierarchically, invoking a lightweight structure-preserving network named as Cross-SRN. Experimental results demonstrate the Cross-SRN achieves competitive or superior restoration performances against the state-of-the-art methods with accurate and clear structural details. Moreover, we set a criterion to select images with rich structural textures. The proposed Cross-SRN outperforms the state-of-the-art methods on the selected benchmark, which demonstrates that our network has a significant advantage in preserving edges.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
47  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
50  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:000835828500005  
Indexed Date:  
2022-08-13  
Institution Address:  
Dalian Univ Technol, Sch Software, Dalian 116620, Peoples R China  
Institution Address:  
Dalian Univ Technol, Int Sch Informat Sci & Engn, 116620, Dalian, Peoples R China  
Institution Address:  
Peking Univ, Sch Elect Engn & Comp Sci, Inst Digital Media, Beijing 100871, Peoples R China  
Institution Address:  
Peng Cheng Lab, Shenzhen 518000, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 52. MFPIDet: improved YOLOV7 architecture based on multi-scale feature

摘要: Prohibited item detection is crucial for the safety of public places. Deep learning, one of the mainstream methods in prohibited item detection tasks, has shown superior performance far beyond traditional prohibited item detection methods. However, most neural network architectures in deep learning still lack sufficient local feature representation ability for overlapping and small targets, and ignore the problem of semantic conflicts caused by direct feature fusion. In this paper, we propose MFPIDet, a novel prohibited item detection neural network architecture based on improved YOLOV7 to achieve reliable prohibited item detection in complex environments. Specifically, a multi-scale attention module (MAM) backbone is proposed to filter the redundant information of target regions and further applied to enhance the local feature representation ability of overlapping objects. Here, to reduce the redundant information of target regions, a squeeze-excitation (SE) block is used to filter the background. Then, aiming at enhancing the feature expression ability of overlapping objects, a multi-scale feature extraction module (MFEM) is designed for local feature representation. In addition, to obtain richer context information, We design an adaptive fusion feature pyramid network (AF-FPN) to combine the adaptive context information fusion module (ACIFM) with the feature fusion module (FFM) to improve the neck structure of YOLOV7. The proposed method is validated on the PIDray dataset, and the tested results showed that our method obtained the highest mAP (68.7%), which is improved by 3.5% than YOLOV7 methods. Our approach provides a new design pattern for prohibited item detection in complex environments and shows the development potential of deep learning in related fields.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
2199-4536  
eISSN:  
2198-6053  
入藏号:  
WOS:001290688300006  
Indexed Date:  
2024-08-18  
Institution Address:  
Chengdu Univ Informat Technol, Sch Comp, Chengdu 610225, Sichuan, Peoples R China  
Institution Address:  
Xihua Univ, Sch Comp & Software Engn, Chengdu 610039, Sichuan, Peoples R China  
Institution:  
Institution:

### 53. MFANet: Multi-scale feature fusion network with attention mechanism

摘要: In order to improve the detection accuracy of the network, it proposes multi-scale feature fusion and attention mechanism net (MFANet) based on deep learning, which integrates pyramid module and channel attention mechanism effectively. Pyramid module is designed for feature fusion in the channel and space dimensions. Channel attention mechanism obtains feature maps in different receptive fields, which divides each feature map into two groups and uses different convolutions to obtain weights. Experimental results show that our strategy boosts state-of-the-arts by 1-2% box AP on object detection benchmarks. Among them, the accuracy of MFANet reaches 34.2% in box AP on COCO dataset. Compared with the current typical algorithms, the proposed method achieves significant performance in detection accuracy.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:000793634200001  
Indexed Date:  
2022-05-22  
Institution Address:  
Hubei Univ Technol, Sch Elect & Elect Engn, Wuhan 430068, Peoples R China  
Institution Address:  
Hubei Univ Technol, Hubei Key Lab High Efficiency Utilizat Solar Ener, Wuhan 430068, Peoples R China  
Institution:  
Institution:

### 54. MSA-Net: Multi-scale feature fusion network with enhanced attention

摘要: Accurate 3D medical imaging can effectively assist doctors in diagnosing diseases. Currently, deep learning-based segmentation methods have yielded good results but face challenges in addressing the complexities of boundary delineation. To address this, we propose an attention- enhanced multi-scale feature fusion network (MSA-Net). This network uses a lightweight module to detect segmentation target locations and then refines the edges of these targets. The MSA-Net mainly consists of an enhanced attention module, a multi-scale feature fusion module, and a cross-layer connectivity module. These modules guide the model to focus on regions of interest, enhancing its context-awareness and ability to capture the intricate and diverse structures in medical images. Ablation experiments demonstrate the effectiveness and significance of each module. We experiment on multiple publicly available datasets, and the results of extensive experiments show that MSA-Net outperforms state-of-the-art techniques.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0045-7906  
eISSN:  
1879-0755  
入藏号:  
WOS:001311886100001  
Indexed Date:  
2024-09-18  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao 266590, Peoples R China  
Institution:

### 55. Multi-Scale Feature Fusion Saliency Object Detection Based on RGB-D

摘要: Purpose salient object detection is a basic problem in computer vision.At present,many saliency detection methods based on deep learning are based on the feature fusion of RGB images and depth maps according to the method of input fusion or result fusion,but these methods cannot effectively fuse of feature maps.In order to improve the performance of salient object detection algorithms,a multi-scale feature fusion RGB-D image salient object detection method is proposed.The main body of the model is designed as two feature encoders,two feature decoders and a cross-model multi-scale feature interleaved fusion module.The two feature encoders correspond to the RGB image and the depth image respectively,which use the ResNet50 network pre-trained by the ImageNet dataset,the feature decoder is used to decode the output of the encoder in 5 different scales,and the cross-model multi-scale feature interleaved fusion module is used for the feature maps of different scales extracted by the decoder and encoder are fused,and the five-level fusion results are spliced and dimensionally reduced to output the final saliency prediction map.Experiments are compared with ten representative models in the past on four public significance data sets.Compared with the second-performing model,the S-measure of the model in this paper is increased by 0.391% on average on each data set.,MAE is decreased by 0.330% on average,and F-measure is decreased by 0.405% on average.A multi-scale feature fusion model is proposed,which abandons the previous fusion method and uses feature fusion to interleave the shallow and deep features.Experiments show that the method proposed in this paper has stronger performance than previous methods,to achieve better results.  
摘要:  
显著性目标检测是计算机视觉的一个基础问题,目前很多基于深度学习的显著性检测方法都是将RGB图像和深度图按照输入融合或结果融合的方法进行特征融合,但这些方法并不能有效地融合特征图,为了提升显著性目标检测算法性能,提出了一种多尺度特征融合的RGB-D图像显著性目标检测方法。将模型主体设计为两个特征编码器、两个特征解码器和一个跨模特多尺度特征交错融合模块。两个特征编码器分别对应RGB图和深度图,其采用经过ImageNet数据集预训练的ResNet50网络,特征解码器用于解码编码器的五种不同尺度的输出,跨模态多尺度特征交错融合模块用于融合解码器和编码器提取的不同尺度的特征图,并将五个层次的融合结果进行拼接和降维,输出最终的显著性预测图。实验在四个公开的显著性数据集上与以往具有代表性的十个模型进行了比较,该模型在各个数据集上,相比于性能第二的模型,S-measure平均提高了0.391%,MAE平均减少了0.330%,F-measure平均减少了0.405%。提出了一种多尺度特征融合模型,摒弃了以往融合的方式,采用特征融合,将浅层和深层的特征分别进行交错融合,实验表明,提出的方法较以往的方法有更强的性能,能够取得更好的效果。  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7753128  
Indexed Date:  
2024-08-30  
Institution Address:  
上海应用技术大学计算机科学与信息工程学院, 上海 201418, 中国  
Institution Address:  
School of Computer Science and Information Engineering,Shanghai Institute of Technology, Shanghai 201418, China  
Institution:  
上海应用技术大学计算机科学与信息工程学院  
Institution:  
School of Computer Science and Information Engineering,Shanghai Institute of Technology

### 56. A Multi-Branch Deep Feature Fusion Network with SAE for Rare Earth

摘要: The Rare Earth Extraction Process (REEP) model is difficult to accurately establish via the extraction mechanism method due to its high complexity. This paper proposes a multi-branch deep feature fusion network with SAE (SAE-MBDFFN) for modeling REEP. We first design a neural network with a multi-branch output structure to simulate the cascade REEP by introducing a multiscale feature fusion mechanism, which can simultaneously concatenate hidden features, original features, and inter-branch coupling features. In order to deal with insufficient labeled data during model training, we then adopt a stacked Sparse Auto-Encoder (SAE) technology to extract the hidden information of mass unlabeled data based on unsupervised learning. This technology can determine the initial parameters of SAE-MBDFFN by unsupervised pretraining. The design methodology of the network is well-founded. Experiments on industrial data indicate that the proposed method has the lowest initial loss value and a faster convergence rate in the fine-tuning stage than other comparison methods, while the prediction accuracy is better well. These results show the effectiveness of the proposed method.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2227-9717  
入藏号:  
WOS:001384020800001  
Indexed Date:  
2024-12-31  
Institution Address:  
East China Jiaotong Univ, Sch Elect & Automat Engn, Nanchang 330013, Peoples R China  
Institution Address:  
East China Jiaotong Univ, State Key Lab Performance Monitoring & Protecting, Nanchang 330013, Peoples R China  
Institution:  
Institution:

### 57. Multi Category Semantic Segmentation of High-resolution Remote Sensing

摘要: Given the imbalance between classes and low classification of multi-class objects in high-resolution remote sensing images,the semantic segmentation method has weak robustness and low accuracy.Based on the convolution neural network model architecture of 2Dmulti-scale feature fusion of UNet,FPN(feature pyramid networks),DeepLabV3+,and HRNet,this paper discusses the multi-scale feature fusion technology and tests and experiments on the model ability of semantic segmentation of fine land cover(up to 16semantic categories)through data preprocessing,loss function,model pre-training,and other factors.Among them,FPN has the highest degree of refinement in semantic segmentation.The FPN model based on efficientbet 1 is pre-trained,the focal loss function is used,the optimal multi-scale feature fusion is selected,and the optimal value of the loss function parameters is optimized and integrated through the super parameter search.Finally,the FPN model with excellent performance is trained.The accuracy is improved by 1.5%,and Kappa is improved by 1.9%,which further improves the models recognition ability and generalization ability for multi analog ground objects.  
摘要:  
针对高分辨遥感影像多类别地物存在类间不平衡、类别区分度低造成的语义分割方法鲁棒性弱和分割精度不高问题,基于UNet、特征金字塔网络(feature pyramid networks,FPN)、DeepLabV3+、HRNet 4种2D多尺度特征融合的卷积神经网络模型架构,对多尺度特征融合技术进行了探讨,通过数据预处理、损失函数、模型预训练等因素对精细土地覆盖的语义分割(多达16个语义类别)进行了模型能力的测试和实验,其中FPN语义分割精细化程度最高。在此基础上,基于EfficientNetB1的FPN模型进行预训练,利用focal loss损失函数选择最优多尺度特征融合,通过超参数搜索损失函数参数最优值进行优化集成,最终训练出性能优异的FPN模型,准确率提高了1.5%,Kappa提高了1.9%,进一步提高了模型对多类别地物的识别能力和泛化能力。  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1000-3177  
入藏号:  
CSCD:7336873  
Indexed Date:  
2023-03-03  
Institution Address:  
河南工业大学信息科学与工程学院, 郑州, 河南 450001, 中国  
Institution Address:  
College of Information Science and Engineering,Henan University of Technology, Zhengzhou, Henan 450001, China  
Institution:  
河南工业大学信息科学与工程学院  
Institution:  
College of Information Science and Engineering,Henan University of Technology

### 58. A Shallow-to-Deep Feature Fusion Network for VHR Remote Sensing Image

摘要: With more detailed spatial information being represented in very-high-resolution (VHR) remote sensing images, stringent requirements are imposed on accurate image classification. Due to the diverse land objects with intraclass variation and interclass similarity, efficient and fine classification of VHR images especially in complex scenes are challenging. Even for some popular deep learning (DL) frameworks, geometric details of land objects may be lost in deep feature levels, so it is difficult to maintain the highly detailed spatial information (e.g., edges, small objects) only relying on the last high-level layer. Moreover, many of the newly developed DL methods require massive well-labeled samples, which inevitably deteriorates the model generalization ability under the few-shot learning. Therefore, in this article, a lightweight shallow-to-deep feature fusion network ((SDFN)-N-2) is proposed for VHR image classification, where the traditional machine learning (ML) and DL schemes are integrated to learn rich and representative information to improve the classification accuracy. In particular, the shallow spectral-spatial features are first extracted and then a novel triple-stage fusion (TSF) module is designed to learn the saliency and discriminative information at different levels for classification. The TSF module includes three feature fusion stages, that is, low-level spectral-spatial feature fusion, middle-level multiscale feature fusion, and high-level multilayer feature fusion. The proposed (SDFN)-N-2 takes the advantage of the shallow-to-deep features, which can extract representative and complementary information from crossing layers. It is important to note that even with limited training samples, the (SDFN)-N-2 still can achieve satisfying classification performance. Experimental results obtained on three real VHR remote sensing datasets including two multispectral and one airborne hyperspectral images covering complex urban scenarios confirm the effectiveness of the proposed approach compared with the state-of-the-art methods.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
24  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
24  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000809416400012  
Indexed Date:  
2022-06-26  
Institution Address:  
Tongji Univ, Coll Surveying & Geoinformat, Shanghai 200092, Peoples R China  
Institution Address:  
Mississippi State Univ, Dept Elect & Comp Engn, Starkville, MS 39762 USA  
Institution Address:  
Univ Trento, Dept Informat Engn & Comp Sci, I-38123 Trento, Italy  
Institution Address:  
Chinese Acad Sci, Xinjiang Inst Ecol & Geog, Urumqi 830011, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 59. AMFF-Net: An attention-based multi-scale feature fusion network for

摘要: Automatic pollen detection based on light microscope (LM) images is helpful for pollinosis symptoms prevention. Recently, many deep learning methods have been proposed to identify pollen grains based on multi-scale feature fusion mechanism. However, in real scenarios, there are two main challenges that need to be considered: (1) Complex pollen characteristics; (2) Irrelevant objects interference. It means that the pollen detection requires not only learning the relationship among multi-scale features but also refining the feature representation. To this end, this paper proposes an attention-based multi-scale feature fusion network (AMFF-Net) for automatic pollen detection on real-world LM images. The proposed AMFF-Net includes three modules: The feature extraction module utilizes the series-connection attention to capture the spatial and channel dependencies of different level feature maps (for solving the challenge 1). In the feature fusion module, a parallel connection attention is able to learn more discriminative feature representation based on bidirectional pathway guidance (for solving the challenge 2). Both are jointly adopted to enhance the representational capacity of the final results in the pollen prediction module. Extensive experiments are conducted on the real-world RPD dataset, and our AMFF-Net achieves the best performance (83.9% of mean average precision) comparing with other state-of-the-art methods. We believe that this work can serve as an important reference for the development of pollen monitoring system in a real scenario.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001069461300001  
Indexed Date:  
2023-10-03  
Institution Address:  
Beijing Univ Technol, Fac Informat Technol, Beijing Engn Res Ctr IoT Software & Syst, Beijing 100124, Peoples R China  
Institution:

### 60. IrisMarkNet: Iris feature watermarking embedding and extraction network

摘要: Based on the urgent need for copyright protection in the digital era, the accuracy of iris recognition technology for authentication must be considered. Ensuring the accuracy of high-precision authentication has tremendous challenges in improving imperceptibility and robustness, especially the longer iris features, as watermarks lead to reduced imperceptibility. A novel digital watermarking method called IrisMarkNet, which embeds the copyright owner's binary iris features into the cover image based on deep neural networks, is first proposed to protect image copyright. It utilizes a novel pyramid feature fusion module PEE based on a multiscale feature fusion strategy to obtain better imperceptibility and well-enhanced robustness, which performs better than other watermark algorithms adopting single-scale feature fusion. Additionally, for different mini-batches, the noise in the noise layer is randomly selected for adversarial training to advance the robustness of the proposed model. In addition, we suggest utilizing Convolutional Block Attention Module (CBAM) Woo et al (2018), which can help to learn better iris features in the decoding stage and propose a novel authenticator to achieve authentication of the image copyright owner. The extensive experimental and comparative results have demonstrated the superior performance of the proposed scheme compared with the state-of-the-art watermark algorithms. Under all experimental distortions, such as JPEG compression, crop attack, Gaussian filter, salt-and-pepper noise, Gaussian noise, and median filter, IrisMarkNet realizes well-improved robustness and imperceptibility along with a good accuracy rate in the authentication of digital images.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000840283900002  
Indexed Date:  
2022-08-23  
Institution Address:  
Shanghai Univ Elect Power, Dept Elect & Informat Engn, Shanghai 200120, Peoples R China  
Institution:

### 61. Emotion Recognition via Multiscale Feature Fusion Network and Attention

摘要: Traditional manual feature-based machine learning methods and deep learning networks have been used for electroencephalogram (EEG)-based emotion recognition in recent years. However, some existing studies ignore the low signal-to-noise ratio and the fact that each subject has unique EEG traits, which suffer from low recognition accuracy and poor robustness. To solve these problems, we propose a novel attention mechanism-based multiscale feature fusion network (AM-MSFFN) that considers high-level features at different scales to improve the generalization of the model for different subjects. Specifically, we first utilize a spatial-temporal convolutional block to extract temporal and spatial features of EEG signals sequentially. Subsequently, considering the sampling rate of EEG signals, the multiscale separable convolutions are designed for capturing emotional state-related information, to better combine and output feature mapping relationships. Convolutional module attention mechanism (CBAM) is applied after point-wise convolution, to better handle EEG variations of different subjects and the key information which facilitates classification. In addition, we adopt a preprocessing module based on data augmentation and data alignment to improve the quality of the training samples. Moreover, ablation studies show that the proposed attention mechanism and multiscale separable convolution contribute significant and consistent gain to the performance of our AM-MSFFN model. To verify the effectiveness of the proposed algorithm, we conducted extensive experiments on the DEAP dataset and SEED. The average accuracies achieve 99.479% and 99.297% for arousal and valence, respectively. The results demonstrated the feasibility of the proposed method.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
10  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
12  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:000991857800055  
Indexed Date:  
2023-06-29  
Institution Address:  
Northwestern Polytech Univ, Sch Elect & Informat, Xian 710072, Shaanxi, Peoples R China  
Institution:

### 62. Multiscale Feature Extraction and Fusion of Image and Text in VQA

摘要: The Visual Question Answering (VQA) system is the process of finding useful information from images related to the question to answer the question correctly. It can be widely used in the fields of visual assistance, automated security surveillance, and intelligent interaction between robots and humans. However, the accuracy of VQA has not been ideal, and the main difficulty in its research is that the image features cannot well represent the scene and object information, and the text information cannot be fully represented. This paper used multi-scale feature extraction and fusion methods in the image feature characterization and text information representation sections of the VQA system, respectively to improve its accuracy. Firstly, aiming at the image feature representation problem, multi-scale feature extraction and fusion method were adopted, and the image features output of different network layers were extracted by a pre-trained deep neural network, and the optimal scheme of feature fusion method was found through experiments. Secondly, for the representation of sentences, a multi-scale feature method was introduced to characterize and fuse the word-level, phrase-level, and sentence-level features of sentences. Finally, the VQA model was improved using the multi-scale feature extraction and fusion method. The results show that the addition of multi-scale feature extraction and fusion improves the accuracy of the VQA model.  
Web of Science 核心合集中的 "被引频次":  
173  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
14  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
174  
ISSN:  
1875-6891  
eISSN:  
1875-6883  
入藏号:  
WOS:000967525800001  
Indexed Date:  
2023-04-30  
Institution Address:  
Univ Elect Sci & Technol China, Sch Automat, Chengdu 610054, Peoples R China  
Institution Address:  
Wenzhou Univ Technol, Sch Data Sci & Artificial Intelligence, Wenzhou 325000, Peoples R China  
Institution Address:  
Guizhou Univ, Coll Resource & Environm Engn, Guiyang 550025, Peoples R China  
Institution Address:  
Louisiana State Univ, Dept Geog & Anthropol, Baton Rouge, LA 70803 USA  
Institution:  
Institution:  
Institution:  
Institution:

### 63. Camouflage target detection based on strong semantic information and

摘要: Aiming at the detection difficulties in camouflage target detection, such as the high similarity between the target and its background, serious damage to the edge, and strong concealment of the target, a camouflage target detection algorithm YOLO of camouflage object detection based on strong semantic information and feature fusion is proposed. First, the attention mechanism convolutional block attention module (CBAM) is constructed to highlight the important channel features and target spatial locations to further aggregate rich semantic information from the high-level feature map. Then the atrous spatial pyramid pooling module is constructed to repeatedly sample the multiscale feature maps to expand the receptive field of the neural network, reduce feature sparsity in the process of convolution, and ensure dense features and multiscale contextual semantic information enter the feature fusion module. Finally, the attention skip-connections are constructed based on the CBAM module for fusing the original feature maps extracted by the backbone network to the corresponding detection outputs so as to eliminate the redundant features as well as enrich the target information of the network outputs. In order to fully verify the performance of the proposed algorithm, a camouflage target detection dataset named strong camouflage efficiency target dataset (SCETD) is constructed. Experimental results on SCETD show that the precision and recall of the proposed algorithm achieve 96.1% and 87.1%, respectively. The AP(0.5) and AP(0.5 : 0.95) achieve 92.3% and 54.4%, respectively. The experimental results prove the effectiveness of the proposed method in detecting camouflage targets.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:001172000900006  
Indexed Date:  
2024-03-17  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Minist Ind & Informat Technol, Key Lab Space Photoelect Detect & Percept, Nanjing, Peoples R China  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Astronaut, Nanjing, Peoples R China  
Institution Address:  
Beijing Inst Space Mech & Elect, Beijing, Peoples R China  
Institution:  
Institution:  
Institution:  
Beijing Inst Space Mech & Elect

### 64. Compression of Multiscale Features of FPN with Channel-Wise Reduction

摘要: With the development of deep learning technology and the abundance of sensors, machine vision applications that utilize vast amounts of image/video data are rapidly increasing in the autonomous vehicle, video surveillance and smart city fields. However, achieving a more compact image/video representation and lower latency solutions is challenging for such machine-based applications. Therefore, it is essential to develop a more efficient video coding standard for machine vision applications. Currently, the Moving Picture Experts Group (MPEG) is developing a new standard called video coding for machines (VCM) with two tracks, each mainly dealing with compression of the input image/video (Track 2) and compression of the features extracted from it (Track 1). In this paper, an enhanced multiscale feature compression (E-MSFC) method is proposed to efficiently compress multiscale features generated by a feature pyramid network (FPN), which is the backbone network of machine vision networks specified in the VCM evaluation framework. The proposed E-MSFC reduces the feature channels to be included in a single feature map and compresses the feature map using versatile video coding (VVC), the latest video standard, rather than the single stream feature compression (SSFC) module in the existing MSFC. In addition, the performance of the E-MSFC is further enhanced by adding a bottom-up structure to the multiscale feature fusion (MSFF) module, which performs the channel-wise reduction in the E-MSFC. Experimental results reveal that the proposed E-MSFC significantly outperforms the VCM image anchor with a BD-rate gain of up to 85.94%, which includes an additional gain of 0.96% achieved by the MSFF with the bottom-up structure.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001031029400001  
Indexed Date:  
2023-08-05  
Institution Address:  
Korea Aerosp Univ, Sch Elect & Informat Engn, Goyang 10540, South Korea  
Institution:

### 65. A feature aggregation and feature fusion network for retinal vessel

摘要: Neural networks have achieved outstanding performance in retinal vessel segmentation. However, since its continuous upsampling and convolution operation in the decoding stage, the semantic information and class information of the high-level features are destroyed. To address these problems, we proposed a new feature aggregation and feature fusion network (FAF-Net). Firstly, we introduced a multi-scale feature aggregation (MFA) block, which adjusts the receptive fields to learn more multi-scale features information. Furthermore, a feature reuse and distribution (FRD) block is intended to preserve the multi-scale feature information of the image and reduce the background noises in the feature map. Finally, the attention feature fusion (AFF) block is employed to effectively reduce the information loss of high-level features and connect the encoding and decoding stages. This multi-path combination helps to learn better representations and more accurate vessel feature maps. We evaluate the network on three retinal image databases (DRIVE, CHASEDB1, STARE). The proposed network outperforms existing current state-of-the-art vessel segmentation methods. Comprehensive experiments prove that FAF-Net is suited to processing medical image segmentation with limited samples and complicated features.  
Web of Science 核心合集中的 "被引频次":  
16  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
18  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:000954681000001  
Indexed Date:  
2023-04-08  
Institution Address:  
HoHai Univ, Coll Internet Things Engn, Changzhou, Peoples R China  
Institution Address:  
Nanjing Marine Radar Inst, Nanjing, Peoples R China  
Institution:  
Institution:  
Nanjing Marine Radar Inst

### 66. Multi-Scale Feature Fusion and Distribution Similarity Network for

摘要: Automatic modulation classification (AMC), as a key technology of cognitive radio, has become a focal point of research. However, most deep learning-based AMC methods require an extensive number of labeled signals to acquire a comprehensive understanding of modulation types, placing substantial pressure on signal acquisition and labeling. To solve this issue, we propose a few-shot AMC (FSAMC) method to facilitate rapid generalization and recognition with limited data, namely multi-scale feature fusion and distribution similarity network (MS2F-DS). Firstly, we design a multi-scale feature fusion (MS2F) model, which aims to extract features with varying fields of view and boost feature fusion, enabling the derivation of contextual information from the signal. Furthermore, we introduce a distribution similarity (DS) classifier to address the insufficient measurement of current similarity measurement functions by considering both micro and macro perspectives of vectors, further increasing intra-class compactness and inter-class separability. Finally, extensive experiments were conducted on 3-way 1, 3, and 5-shot FSAMC tasks using public datasets RML2016.10a and RML2016.10b, and the results demonstrated the effectiveness of our method.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1070-9908  
eISSN:  
1558-2361  
入藏号:  
WOS:001342549800002  
Indexed Date:  
2024-11-07  
Institution Address:  
Xidian Univ, Xian 710126, Peoples R China  
Institution:

### 67. MSB-Net: An End-to-End Network for Extracting Building from

摘要: Extracting buildings from high-resolution remote sensing imagery (HRSI) is of great significance to emergency management, land resource utilization, and analysis, as well as city planning and construction. However, due to the complex backgrounds and diverse appearances and different sizes of buildings in HRSI, most existing methods for automatic building extraction are difficult to obtain strong building feature representation from low-level and high-level features. Furthermore, existing research mainly focused on regional accuracy, whereas less attention was paid to the description of building boundaries. In this article, MSB-Net, an end-to-end neural network, is proposed to address these issues. A multiscale feature fusion module (MSFFM) is designed to capture and fuse multiscale features. A local branch (LB) constructed by the MSFFM and position attention, is used to obtain long range of context information between different positions and extract the essential features of buildings (e.g., shapes, edges) from low-level features. And a global branch (GB) is designed to use the MSFFM and channel attention to enhance high-level features. Therefore, our method can not only obtain information on building-related attribute categories, but also capture the rich context information in channel dimensions. The boundary enhancement and completion module take the output of the GB and LB as input to search for the missing parts and details of buildings to improve the segmentation accuracy and boundary quality. Our method is tested on two public building datasets and achieves superior classification performance.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001231562600006  
Indexed Date:  
2024-06-15  
Institution Address:  
Guilin Univ Technol, Coll Geomat & Geoinformat, Guilin 541006, Peoples R China  
Institution:

### 68. Multiscale Feature Fusion and Graph Convolutional Network for Detecting

摘要: With the emergence of blockchain technology, the cryptocurrency market has experienced significant growth in recent years, simultaneously fostering environments conducive to cybercrimes such as phishing scams. Phishing scams on blockchain platforms like Ethereum have become a grave economic threat. Consequently, there is a pressing demand for effective detection mechanisms for these phishing activities to establish a secure financial transaction environment. However, existing methods typically utilize only the most recent transaction record when constructing features, resulting in the loss of vast amounts of transaction data and failing to adequately reflect the characteristics of nodes. Addressing this need, this study introduces a multiscale feature fusion approach integrated with a graph convolutional network model to detect phishing scams on Ethereum. A node basic feature set comprising 12 features is initially designed based on the Ethereum transaction dataset in the basic feature module. Subsequently, in the edge embedding representation module, all transaction times and amounts between two nodes are sorted, and a gate recurrent unit (GRU) neural network is employed to capture the temporal features within this transaction sequence, generating a fixed-length edge embedding representation from variable-length input. In the time trading feature module, attention weights are allocated to all embedding representations surrounding a node, aggregating the edge embedding representations and structural relationships into the node. Finally, combining basic and time trading features of the node, graph convolutional networks (GCNs), SAGEConv, and graph attention networks (GATs) are utilized to classify phishing nodes. The performance of these three graph convolution-based deep learning models is validated on a real Ethereum phishing scam dataset, demonstrating commendable efficiency. Among these, SAGEConv achieves an F1-score of 0.958, an AUC-ROC value of 0.956, and an AUC-PR value of 0.949, outperforming existing methods and baseline models.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2079-9292  
入藏号:  
WOS:001191478900001  
Indexed Date:  
2024-04-02  
Institution Address:  
Hainan Normal Univ, Coll Informat Sci Technol, 99 Longkun South Rd, Haikou 571158, Peoples R China  
Institution:

### 69. PSRGAN: Generative Adversarial Networks for Precipitation Downscaling

摘要: This study proposes an innovative generative adversarial network (GAN)-based downscaling model for precipitation, named PSRGAN, which aims to enhance the spatial resolution of meteorological data using deep learning techniques. The PSRGAN model integrates a multiscale feature fusion module (Rception), a kernel attention module (KAM), and the generator-discriminator framework of GANs to address challenges such as data sparsity and spatiotemporal correlations that traditional precipitation super-resolution (SR) methods struggle with. By extracting multiscale spatial features, PSRGAN improves the model's ability to detect key precipitation regions and enhances the accuracy of predicting extreme precipitation events. The model is trained and tested using low-resolution (LR) and high-resolution (HR) simulated datasets based on regional climate models (RCMs), with performance evaluated through various metrics. The experimental results demonstrate that PSRGAN achieves strong performance in the precipitation downscaling task.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001432988800010  
Indexed Date:  
2025-03-05  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Elect & Informat Engn, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Comp Sci, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Artificial Intelligence, Nanjing 210044, Peoples R China  
Institution:  
Institution:  
Institution:

### 70. Multiscale and Multidirection Feature Extraction Network for

摘要: Deep learning (DL) plays an increasingly important role in Earth observation by multisource remote sensing. However, the current DL-based methods do not make a fully use of the complementary information among multisource remote sensing data, such as hyperspectral image and light detection and ranging data, and lack the consideration of multiscale, directional, and fine-grained features. To address these issues, a multiscale and multidirection feature extraction network is proposed in this article. Specifically, the multiscale spatial feature (MSSpaF) module is designed to extract the MSSpaFs, and then, these features are fused by feature concatenation operation. In addition, the multidirection spatial feature module is designed to further extract multidirection and frequency information, employing cross-layer connection and multiscale feature fusion strategy to improve the fineness of the proposed network. Moreover, the spectral feature module is employed to provide detailed spectral information for enhancing the expression ability of multiscale features. Experimental results on three different datasets demonstrate the superior classification performance of the proposed framework.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001241740900007  
Indexed Date:  
2024-06-13  
Institution Address:  
Changan Univ, Sch Elect & Control Engn, Xian 710064, Peoples R China  
Institution Address:  
Beijing Inst Technol, Sch Informat & Elect, Beijing 100081, Peoples R China  
Institution:  
Institution:

### 71. Denoising Multiscale Back-Projection Feature Fusion for Underwater Image

摘要: In recent decades, enhancing underwater images has become a crucial challenge when obtaining high-quality visual information in underwater environment detection, attracting increasing attention. Original underwater images are affected by a variety of underwater environmental factors and exhibit complex degradation phenomena such as low contrast, blurred details, and color distortion. However, most encoder-decoder-based methods fail to restore the details of underwater images due to information loss during downsampling. The noise in images also influences the recovery of underwater images with complex degradation. In order to address these challenges, this paper introduces a simple but effective denoising multiscale back-projection feature fusion network, which represents a novel approach to restoring underwater images with complex degradation. The proposed method incorporates a multiscale back-projection feature fusion mechanism and a denoising block to restore underwater images. Furthermore, we designed a multiple degradation knowledge distillation strategy to extend our method to enhance various types of degraded images, such as snowy images and hazy images. Extensive experiments on the standard datasets demonstrate the superior performance of the proposed method. Qualitative and quantitative analyses validate the effectiveness of the model compared to several state-of-the-art models. The proposed method outperforms previous deep learning models in recovering both the blur and color bias of underwater images.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
eISSN:  
2076-3417  
入藏号:  
WOS:001246685900001  
Indexed Date:  
2024-06-18  
Institution Address:  
Dalian Maritime Univ, Comp Sci & Technol, Dalian 116026, Peoples R China  
Institution:

### 72. Res-MulFra: Multilevel and Multiscale Framework for Brain Tumor

摘要: In clinical diagnosis and surgical planning, extracting brain tumors from magnetic resonance images (MRI) is very important. Nevertheless, considering the high variability and imbalance of the brain tumor datasets, the way of designing a deep neural network for accurately segmenting the brain tumor still challenges the researchers. Moreover, as the number of convolutional layers increases, the deep feature maps cannot provide fine-grained spatial information, and this feature information is useful for segmenting brain tumors from the MRI. Aiming to solve this problem, a brain tumor segmenting method of residual multilevel and multiscale framework (Res-MulFra) is proposed in this article. In the proposed framework, the multilevel is realized by stacking the proposed RMFM-based segmentation network (RMFMSegNet), which is mainly used to leverage the prior knowledge to gain a better brain tumor segmentation performance. The multiscale is implemented by the proposed RMFMSegNet, which includes both the parallel multibranch structure and the serial multibranch structure, and is mainly designed for obtaining the multiscale feature information. Moreover, from various receptive fields, a residual multiscale feature fusion module (RMFM) is also proposed to effectively combine the contextual feature information. Furthermore, in order to gain a better brain tumor segmentation performance, the channel attention module is also adopted. Through assessing the devised framework on the BraTS dataset and comparing it with other advanced methods, the effectiveness of the Res-MulFra is verified by the extensive experimental results. For the BraTS2015 testing dataset, the Dice value of the proposed method is 0.85 for the complete area, 0.72 for the core area, and 0.62 for the enhanced area.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0899-9457  
eISSN:  
1098-1098  
入藏号:  
WOS:001262983200001  
Indexed Date:  
2024-07-20  
Institution Address:  
Natl Key Lab Sci & Technol Commun, Chengdu, Peoples R China  
Institution Address:  
Univ Elect Sci & Technol China, Sch Informat & Software Engn, Network & Data Secur Key Lab Sichuan Prov, Chengdu, Peoples R China  
Institution Address:  
YIBIN GREAT Technol Co Ltd, Yibin, Peoples R China  
Institution Address:  
Ningbo WebKing Technol Joint Stock Co Ltd, Ningbo, Peoples R China  
Institution:  
Natl Key Lab Sci & Technol Commun  
Institution:  
Institution:  
YIBIN GREAT Technol Co Ltd  
Institution:  
Ningbo WebKing Technol Joint Stock Co Ltd

### 73. Stepwise Attention-Guided Multiscale Fusion Network for Lightweight and

摘要: Many exceptional deep learning networks have demonstrated remarkable proficiency in general object detection tasks. However, the challenge of detecting ships in synthetic aperture radar (SAR) imagery increases due to the complex and various nature of these scenes. Moreover, sophisticated large-scale models necessitate substantial computational resources and hardware expenses. To address these issues, a new framework is proposed called a stepwise attention-guided multiscale feature fusion network (SAFN). Specifically, we introduce a stepwise attention mechanism designed to selectively emphasize relevant information and filter out irrelevant details of objects in a step-by-step manner. Firstly, a novel LGA-FasterNet is proposed, which incorporates a lightweight backbone FasterNet with lightweight global attention (LGA) to realize expressive feature extraction while reducing the model's parameters. To effectively mitigate the impact of scale and complex background variations, a deformable attention bidirectional fusion network (DA-BFNet) is proposed, which introduces a novel deformable location attention (DLA) block and a novel deformable recognition attention (DRA) block, strategically integrating through bidirectional connections to achieve enhanced features fusion. Finally, we have substantiated the robustness of the new framework through extensive testing on the publicly accessible SAR datasets, HRSID and SSDD. The experimental outcomes demonstrate the competitive performance of our approach, showing a significant enhancement in ship detection accuracy compared to some state-of-the-art methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001311533900001  
Indexed Date:  
2024-09-16  
Institution Address:  
Shanghai Univ Engn Sci, Sch Elect & Elect Engn, Shanghai 201602, Peoples R China  
Institution Address:  
Shanghai Inst Satellite Engn, Shanghai 200240, Peoples R China  
Institution Address:  
Dongshin Univ, Dept Comp Engn, Naju 58245, South Korea  
Institution Address:  
Harbin Inst Technol, Sch Elect & Informat Engn, Harbin 150001, Peoples R China  
Institution:  
Institution:  
Shanghai Inst Satellite Engn  
Institution:  
Institution:

### 74. Integration of multiscale fusion of residual neural network with 2-D

摘要: The human lower limb movements recognition (LLMR) plays a pivotal role in active lower limb exoskeleton robots. Employing surface electromyography (sEMG) signals for LLMR allows for the convenient, rapid and stable capture of signal variations, facilitating efficient identification of lower limb motion patterns. However, current sEMG-based LLMR methods face challenges such as incomplete feature extraction, limited contextual information and restricted feature extraction scales during feature extraction. This paper proposed a LLMR method based on Gramian Angular Fields (GAF) and multiscale fusion of Residual Neural Network (MS-ResNet). The denoised sEMG time series was transformed into Gramian Angular Difference Field (GADF) matrix based on GAF. The MS-ResNet model, incorporating ResNet and multiscale feature fusion concepts, was proposed to comprehensively capture global and local information through different-scale feature extraction and fusion, so as to improve recognition performance. sEMG signals from 11 muscles of the preferred leg of 15 healthy subjects were recorded during six common lower limb movements. Experimental analysis investigated the impact of the convolutional kernel size (k k x k ) in Stream 2 of MS-ResNet and the number of muscles involved on recognition performance. The study revealed that selecting k as 13, coupled with 11 muscles, yielded optimal model performance with the average cross-individual recognition accuracy reaching 97.62 %, demonstrating the model's efficiency in LLMR. This method could provide a viable solution for developing more efficient and reliable LLMR systems, applicable to lower limb exoskeleton robots and intelligent prosthetics.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:001333485000001  
Indexed Date:  
2024-10-21  
Institution Address:  
Chinese Acad Sci, Shenzhen Inst Adv Technol, Shenzhen 518000, Peoples R China  
Institution Address:  
Southern Univ Sci & Technol, Shenzhen 518000, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Hong Kong 999077, Peoples R China  
Institution:  
Institution:  
Institution:

### 75. Multi-Scale Fusion Siamese Network Based on Three-Branch Attention

摘要: Remote sensing image change detection (CD) is an important means in remote sensing data analysis tasks, which can help us understand the surface changes in high-resolution (HR) remote sensing images. Traditional pixel-based and object-based methods are only suitable for low- and medium-resolution images, and are still challenging for complex texture features and detailed image detail processing in HR images. At present, the method based on deep learning has problems such as inconsistent fusion and difficult model training in the combination of the difference feature information of the deep and shallow layers and the attention mechanism, which leads to errors in the distinction between the changing region and the invariant region, edge detection and small target detection. In order to solve the above problems of inconsistent fusions of feature information aggregation and attention mechanisms, and indistinguishable change areas, we propose a multi-scale feature fusion Siamese network based on attention mechanism (ABMFNet). To tackle the issues of inconsistent fusion and alignment difficulties when integrating multi-scale fusion and attention mechanisms, we introduce the attention-based multi-scale feature fusion module (AMFFM). This module not only addresses insufficient feature fusion and connection between different-scale feature layers, but also enables the model to automatically learn and prioritize important features or regions in the image. Additionally, we design the cross-scale fusion module (CFM) and the difference feature enhancement pyramid structure (DEFPN) to assist the AMFFM module in integrating differential information effectively. These modules bridge the spatial disparity between low-level and high-level features, ensuring efficient connection and fusion of spatial difference information. Furthermore, we enhance the representation and inference speed of the feature pyramid by incorporating a feature enhancement module (FEM) into DEFPN. Finally, the BICD dataset proposed by the laboratory and public datasets LEVIR-CD and BCDD are compared and tested. We use F1 score and MIoU values as evaluation metrics. For AMBMFNet, the F1 scores on the three datasets are 77.69%, 81.57%, and 77.91%, respectively, while the MIoU values are 84.65%, 85.84%, and 84.54%, respectively. The experimental results show that ABMFNet has better effectiveness and robustness.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
7  
eISSN:  
2072-4292  
入藏号:  
WOS:001231601000001  
Indexed Date:  
2024-05-31  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Collaborat Innovat Ctr Atmospher Environm & Equipm, B DAT, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Forestry Univ, Coll Informat Sci & Technol, Nanjing 210037, Peoples R China  
Institution:  
Institution:

### 76. Vehicle color recognition based on smooth modulation neural network with

摘要: Vehicle Color Recognition (VCR) plays a vital role in intelligent traffic management and criminal investigation assistance. However, the existing vehicle color datasets only cover 13 classes, which can not meet the current actual demand. Besides, although lots of efforts are devoted to VCR, they suffer from the problem of class imbalance in datasets. To address these challenges, in this paper, we propose a novel VCR method based on Smooth Modulation Neural Network with Multi-Scale Feature Fusion (SMNN-MSFF). Specifically, to construct the benchmark of model training and evaluation, we first present a new VCR dataset with 24 vehicle classes, Vehicle Color-24, consisting of 10091 vehicle images from a 100-hour urban road surveillance video. Then, to tackle the problem of long-tail distribution and improve the recognition performance, we propose the SMNN-MSFF model with multiscale feature fusion and smooth modulation. The former aims to extract feature information from local to global, and the latter could increase the loss of the images of tail class instances for training with class-imbalance. Finally, comprehensive experimental evaluation on Vehicle Color-24 and previously three representative datasets demonstrate that our proposed SMNN-MSFF outperformed state-of-the-art VCR methods. And extensive ablation studies also demonstrate that each module of our method is effective, especially, the smooth modulation efficiently help feature learning of the minority or tail classes. Vehicle Color-24 and the code of SMNN-MSFF are publicly available and can contact the author to obtain.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2095-2228  
入藏号:  
CSCD:7477509  
Indexed Date:  
2023-09-01  
Institution Address:  
School of Communications and Information Engineering & School of Artificial Intelligence, Xian University of Posts & Telecommunications, Xi'an, Shaanxi 710121, China  
Institution Address:  
School of Computer Science and Technology, University of Science and Technology of China, Hefei, Anhui 230026, China  
Institution:  
School of Communications and Information Engineering & School of Artificial Intelligence, Xian University of Posts & Telecommunications  
Institution:  
School of Computer Science and Technology, University of Science and Technology of China

### 77. Multiscale feature fusion and cross-guidance for few-shot semantic

摘要: Objective Few-shot semantic segmentation is one of the fundamental and challenging tasks in the field of computer vision. It aims to use a limited amount of annotated support samples to guide the segmentation of unknown objects in a query image. Compared with traditional semantic segmentation, few-shot semantic segmentation methods effectively alleviate problems, such as the high cost of per-pixel annotation greatly limiting the application of semantic segmentation technology in practical scenarios and the weak generalization ability of this model for novel class targets. The existing few-shot semantic segmentation methods mainly utilize the meta-learning architecture with dual-branch networks, where the support branch consists of the support images and their corresponding per-pixel labeled ground truth masks, and the query branch takes the input of the new image to be segmented, and both branches share the same semantic classes. The valuable information of support images in the support branch is extracted to guide the segmentation of unknown novel classes in query images. However, different instances of the same semantic class may have variations in appearance and scale, and the information extracted solely from the support branch is insufficient to guide the segmentation of unknown novel classes in query images. Although some researchers have attempted to improve the performance of few-shot semantic segmentation through bidirectional guidance, existing bidirectional guidance models overly rely on the pseudo masks predicted by the query branch in the intermediate stage. If the initial predictions of the query branch are poor, it can easily lead to a weak generalization of shared semantics, which is not conducive to improving segmentation performance. Method A multiscale feature fusion and cross-guidance network for few-shot semantic segmentation is proposed to alleviate these problems, attempting to construct the information interaction between the support branch and the query branch to improve the performance of the few-shot semantic segmentation task. First, a set of pretrained backbone networks with shared weights are used as feature extractors to map features from the support and query branch into the same deep feature space, and then the low-level, intermediate-level, and high-level features output by them are fused at multiple scales to construct a multiscale feature set, which enriches the semantic information of features and enhances the reliability of the feature expression. Second, with the help of the ground-truth mask of the support branch, the fused support features are decomposed into the target-related foreground feature maps and task-irrelevant background feature maps. Then, a feature interaction module is designed on the basis of the cross-attention mechanism, which establishes information interaction between the target-related foreground feature maps of the support branch and the entire query branch feature map, aiming to promote the interactivity between branches while enhancing the expressiveness of task-related features. In addition, a mask average pooling strategy is used on the interactive feature map to generate a target foreground region prototype set, and a background prototype set is generated on the support background feature map. Finally, the cosine similarity measure is used to calculate the similarity values between the support features and the prototype sets and between the query features and the prototype sets; then, the corresponding mask is generated on the basis of the maximum similarity value at each position.  
摘要:  
目的构建支持分支和查询分支间的信息交互对于提升小样本语义分割的性能具有重要作用,提出一种多尺度特征融合与交叉指导的小样本语义分割算法。方法利用一组共享权重的主干网络将双分支输入图像映射到深度特征空间,并将输出的低层、中间层和高层特征进行尺度融合,构造多尺度特征;借助支持分支的掩码将支持特征分解成目标前景和背景特征图;设计了一种特征交互模块,在支持分支的目标前景和整个查询分支的特征图上建立信息交互,增强任务相关特征的表达能力,并利用掩码平均池化策略生成目标前景和背景区域的原型集;利用无参数的度量方法分别计算支持特征和原型集、查询特征与原型集之间的余弦相似度值,并根据相似度值给出对应图像的掩码。结果通过在PASCAL-5i (pattern analysis, statistical modeling and computational learning)和COCO-20i (common objects in context)开源数据集上进行实验,结果表明,利用VGG-16(Visual Geometry Group)、ResNet-50 (residual neural network)和ResNet-101作为主干网络时,所提模型在1-way 1-shot任务中,分别获得50.2%、53.2%、 57.1%和23.9%、35.1%、36.4%的平均交并比(mean intersection over union,mIoU),68.3%、69.4%、72.3%/和60.1%、62.4%、64.1%的前景背景二分类交并比(foreground and background intersection over union,FB-IoU);在1-way 5-shot任务上,分别获得52.9%、55.7%、59.7%和32.5%、37.3%、38.3%的mIoU,69.7%、72.5%、74.6%和64.2%、66.2%、66.7%的FB-IoU。结论相比当前主流的小样本语义分割模型,所提模型在1-way 1-shot和1-way 5-shot任务中可以获得更高的mIoU和FB-IoU,综合性能提升效果显著。  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1006-8961  
入藏号:  
CSCD:7725003  
Indexed Date:  
2024-08-02  
Institution Address:  
晋中职业技术学院电子信息系, 晋中, 山西 030600, 中国  
Institution Address:  
英国布莱顿大学智能和可用系统中心, 布莱顿, BN24AT, 英国  
Institution Address:  
Department of Electronic Information,Jingzhong Vocational and Technical College, Jinzhong, Shanxi 030600, China  
Institution Address:  
Intelligent and Available Systems Centre,University of Brighton, Brighton, BN24AT, England  
Institution:  
晋中职业技术学院电子信息系  
Institution:  
英国布莱顿大学智能和可用系统中心  
Institution:  
Department of Electronic Information,Jingzhong Vocational and Technical College  
Institution:  
Intelligent and Available Systems Centre,University of Brighton

### 78. Identification of zinc stripping defects from cathode plate based on

摘要: During hydro-zinc smelting, the cathode plates are attached by with residual zinc or discarded due to damaged insulation strips and edging strips. Such defects limit the recycling of cathode plates. Current manual observation leads to low accuracy and speed of recognition owing to perception biases. Therefore, this work applied computer vision and deep learning semantic segmentation technology to realize the defect recognition of cathode plates. Firstly, a semantic segmentation dataset on cathode plates was constructed for training and testing the model. Then a network of attention mechanism and multiscale feature fusion (AMNet) was proposed to detect the defects. In AMNet, the encoder-decoder jump connection architecture was designed to fuse low-level and highlevel features. A channel attention module was incorporated to enhance focus on the channels with important information, and the newly proposed multiscale feature extraction module was used to solve the problem of target multiscale capture. Through related parameter selection experiments, the final AMNet achieved 95.12% and 97.73% for Mean Intersection over Union (MIoU) and mean pixel accuracy (MPA), respectively. These values are 3.24 and 1.74 percentage points higher than DeepLabv3+.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001440672300001  
Indexed Date:  
2025-03-16  
Institution Address:  
Cent South Univ, Coll Mech & Elect Engn, Changsha 410083, Peoples R China  
Institution Address:  
Cent South Univ, State Key Lab Precis Mfg Extreme Serv Performance, Changsha 410083, Peoples R China  
Institution Address:  
Hunan CHMM Sunwards New Mat Co Ltd, Changsha 410000, Peoples R China  
Institution:  
Institution:  
Institution:  
Hunan CHMM Sunwards New Mat Co Ltd

### 79. Clustering Feature Constraint Multiscale Attention Network for Shadow

摘要: Shadow extraction is an important and challenging task in remote sensing image analysis because the presence of shadows not only reduces radiation information but also affects the interpretation of remote sensing images. In this article, a clustering feature constraint multiscale attention network for shadow extraction from remote sensing images is proposed. First, in addition to the pixel-level description of the traditional neural network, our method focuses on the clustering relationships between pixel pairs to obtain the pixel group features of shadows. The feature extraction capability of the network is improved with a reweighting mechanism at the pixel level and pixel group features. Second, we employ a feature fusion algorithm by considering contextual information to improve the network's attention toward shadow areas and enhance the nonlinear expression ability during the encoding and decoding layers. Furthermore, considering the most prominent multiscale features of shadows in remote sensing images, a deep multiscale feature aggregation structure is established to better fit the multiscale feature expression of shadows. Finally, we construct a shadow extraction dataset to verify the proposed approach. We compare our method with the results of state-of-the-art deep learning models. The results show that the intersection over union (IOU) of our method is improved by 0.85%-9.51% and that the F1-score is improved by 0.73-6.48. In addition, the test results for images with different resolutions prove that the proposed approach is more robust than the other methods.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
23  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
23  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000773300900024  
Indexed Date:  
2022-04-06  
Institution Address:  
Southwest Jiaotong Univ, Fac Geosci & Environm Engn, Chengdu 611756, Peoples R China  
Institution Address:  
Sejong Univ, Dept Software, Seoul 05006, South Korea  
Institution:  
Institution:

### 80. Detecting Dim Small Target in Infrared Images via Subpixel Sampling

摘要: Infrared dim small target (IDST) detection is regarded as a critical technology for the interpretation of space-based remote sensing images. In recent years, driven by deep learning technology and the surge of data, remarkable effects have been achieved for dim small target detection in infrared images. Nevertheless, the intrinsic feature scarcity and low signal-to-clutter ratio (SCR) characteristics pose tremendous challenges to deep learning-based detection methods. In this letter, we present a novel subpixel sampling cuneate network (SPSCNet) to detect dim small targets in infrared images. The overall model architecture is based on an end-to-end cuneate network with multiple groups of parallel high-to-low resolution subnetworks. Specifically, we design a multiscale feature reweighted fusion (MSFRF) module to effectively fuse multiscale feature maps which contain both low-level detail features and high-level semantics information. In addition, considering that the pooling operation may lose dim small targets with low SCR, we also exploit a subpixel sampling scheme to greatly retain the features of small targets. Moreover, to better test and verify the performance of the proposed method, we also develop an IDST dataset to conduct more comparative experiments. Extensive experiments on the single-frame infrared small target (SIRST) and IDST datasets illustrate that the proposed SPSCNet yields state-of-the-art performance in comparison with other detection algorithms.  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000838529500005  
Indexed Date:  
2022-08-20  
Institution Address:  
Natl Univ Def Technol, Coll Elect Sci & Technol, Changsha 410073, Peoples R China  
Institution:

### 81. MSFF-UNet: Image segmentation in colorectal glands using an

摘要: Glands are closely related to the diagnosis of tumors. In pathological images, segmentation of the colorectal gland is a prerequisite for quantitative diagnosis. Segmentation algorithms based on deep learning have been widely used in medical images. However, the existing segmentation method has feature fusion only existing in adjacent layers, ignoring cross-layer fusion. And ignoring the combination of local and global information for graphics. To solve the above problems, we propose a multi-scale fusion model (MSFF-UNet) based on U-Net. We enhance the fusion of multi-scale information in the feature fusion module (FFM) and combine spatial attention to highlight the spatial structure of objects. In addition, we use the receptive field extension module (RFEM) to fuse local and global information, thereby reducing information loss and improving segmentation performance. We also propose a boundary loss function, which enables the network to pay more attention to the boundary information and make the segmentation results more accurate. Compared to the U-Net model, our network improved the DICE score by 1.95% and the MIOU score by 2.6%, effectively improving the accuracy of colorectal glandular segmentation.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001083978800014  
Indexed Date:  
2023-12-19  
Institution Address:  
Chengdu Univ Technol, Coll Comp Sci & Cyber Secur, Chengdu 610059, Peoples R China  
Institution:

### 82. Pyramid attention object detection network with multi-scale feature

摘要: With the development of deep learning, object detection has made substantial progress. However, when the object to be detected in the image is small or partially occluded, the detection network often fails to detect it successfully. We propose a multi-scale feature fusion pyramid attention module, which effectively combines the global average pooling results of multiple scales with the upper features in the residual blocks of the feature extraction network to obtain more spatial context information in the original feature map. We added the multi-scale feature fusion pyramid attention module proposed in this paper based on YoloV3 and conducted experiments on the PASCALL VOC and MS COCO datasets. The experimental results show that the attention module can effectively help the network detect small objects and accurately detect partially occlusion objects.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
0045-7906  
eISSN:  
1879-0755  
入藏号:  
WOS:000879218900002  
Indexed Date:  
2022-11-18  
Institution Address:  
Yangzhou Univ, Sch Informat Engn, Yangzhou, Peoples R China  
Institution Address:  
Kyushu Inst Technol, Sch Engn, Kitakyushu, Japan  
Institution:  
Institution:

### 83. MFP-Net: Multi-scale feature pyramid network for crowd counting

摘要: Although deep learning has been widely used for dense crowd counting, it still faces two challenges. Firstly, the popular network models are sensitive to scale variance of human head, human occlusions, and complex background due to repeated utilization of vanilla convolution kernels. Secondly, the vanilla feature fusion often depends on summation or concatenation, which ignores the correlation of different features leading to information redundancy and low robustness to background noise. To address these issues, a multi-scale feature pyramid network (MFP-Net) for dense crowd counting is proposed in this paper. The proposed MFP-Net makes two contributions. Firstly, the feature pyramid fusion module is designed that adopts rich convolutions with different depths and scales, not only to expand the receptive field, but also to improve the inference speed of models by using parallel group convolution. Secondly, a feature attention-aware module is added in the feature fusion stage. The module can achieve local and global information fusion by capturing the importance of the spatial and channel domains to improve model robustness. The proposed MFP-Net is evaluated on five publicly available datasets, and experiments show that the MFP-Net not only provides better crowd counting results than comparative models, but also requires fewer parameters.  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
1751-9659  
eISSN:  
1751-9667  
入藏号:  
WOS:000659122700001  
Indexed Date:  
2021-06-25  
Institution Address:  
Shaanxi Univ Sci & Technol, Shaanxi Joint Lab Artificial Intelligence, Xian, Peoples R China  
Institution Address:  
Shaanxi Univ Sci & Technol, Sch Elect Informat & Artificial Intelligence, Xian, Peoples R China  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Automat, Xian 710121, Peoples R China  
Institution Address:  
Shaanxi Univ Sci & Technol, Sch Elect & Control Engn, Xian, Peoples R China  
Institution Address:  
Brunel Univ London, Dept Elect & Elect Engn, Uxbridge, Middx, England  
Institution Address:  
Xi An Jiao Tong Univ, Sch Mech Engn, Xian, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 84. Multi-scale high and low feature fusion attention network for intestinal

摘要: Accurate abnormality classification in intestinal images is critical for the diagnosis and treatment of early-stage intestinal cancers, but remains challenging due to the large intra-lesion variability and high similarity between abnormal lesions and normal tissue. To solve the above problems, a multi-scale high and low feature fusion attention network is proposed to efficiently utilize the features extracted by the backbone network. First, a multi-scale feature extraction module is used to extract features, and then a detail capture attention module and a dense sampling fusion module are used to focus on and fuse the lesion information hidden in the superficial and deep layers. Finally, classification effect is obtained by fusing the lesion information. Experiments show that the method achieves 98% classification accuracy on the Kvasir dataset with three types of normal, polyp and ulcer. In addition, the classification accuracy in the private dataset reaches 97.52% and the experimental effect is better than other existing deep learning classification methods.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:000935498700001  
Indexed Date:  
2023-03-13  
Institution Address:  
Zhejiang Univ Technol, Coll Informat Engn, Hangzhou 310023, Zhejiang, Peoples R China  
Institution Address:  
Lishui Peoples Hosp, Lishui 331100, Zhejiang, Peoples R China  
Institution:  
Institution:  
Lishui Peoples Hosp

### 85. Unmanned Aerial Vehicle Object Detection Based on Information-Preserving

摘要: General deep learning methods achieve high-level semantic feature representation by aggregating hierarchical features, which performs well in object detection tasks. However, issues arise with general deep learning methods in UAV-based remote sensing image object detection tasks. Firstly, general feature aggregation methods such as stride convolution may lead to information loss in input samples. Secondly, common FPN methods introduce conflicting information by directly fusing feature maps from different levels. These shortcomings limit the model's detection performance on small and weak targets in remote sensing images. In response to these concerns, we propose an unmanned aerial vehicle (UAV) object detection algorithm, IF-YOLO. Specifically, our algorithm leverages the Information-Preserving Feature Aggregation (IPFA) module to construct semantic feature representations while preserving the intrinsic features of small objects. Furthermore, to filter out irrelevant information introduced by direct fusion, we introduce the Conflict Information Suppression Feature Fusion Module (CSFM) to improve the feature fusion approach. Additionally, the Fine-Grained Aggregation Feature Pyramid Network (FGAFPN) facilitates interaction between feature maps at different levels, reducing the generation of conflicting information during multi-scale feature fusion. The experimental results on the VisDrone2019 dataset demonstrate that in contrast to the standard YOLOv8-s, our enhanced algorithm achieves a mean average precision (mAP) of 47.3%, with precision and recall rates enhanced by 6.3% and 5.6%, respectively.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
eISSN:  
2072-4292  
入藏号:  
WOS:001277557000001  
Indexed Date:  
2024-08-02  
Institution Address:  
Natl Univ Def Technol NUDT, Natl Key Lab Sci & Technol ATR, Changsha 410073, Peoples R China  
Institution:

### 86. PAF-Net: A Progressive and Adaptive Fusion Network for Pavement Crack

摘要: Automatic crack detection remains challenging due to factors such as irregular crack shapes and sizes, uneven illumination, complex backgrounds, and image noise. Deep learning has shown promise in computer vision for pixel-wise crack detection, but existing methods still suffer from limitations such as information loss, insufficient feature fusion, and semantic gap issues. To address these challenges, a novel pavement crack segmentation network, called PAF-Net, is proposed, which incorporates progressive and adaptive feature fusion. To mitigate information loss caused by feature downsampling, a progressive context fusion (PCF) block is introduced to capture context information from adjacent scales. To better capture strong features from local regions, a dual attention (DA) block is proposed that leverages both global and local context information, reducing the semantic gap issue. Furthermore, to achieve effective multi-scale feature fusion, a dynamic weight learning (DWL) block is proposed that enables efficient fusion of feature maps from different network layers. Additionally, a multi-scale input unit is incorporated to provide the proposed segmentation network with more contextual information. To evaluate the performance of PAF-Net, we conduct experiments using four common evaluation metrics and compare it with multiple mainstream segmentation models on three public datasets. The proposed PAF-Net demonstrates superior segmentation accuracy for pixel-level crack detection compared to other segmentation models, as evident from qualitative and quantitative experimental results.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
17  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
17  
ISSN:  
1524-9050  
eISSN:  
1558-0016  
入藏号:  
WOS:001025570400001  
Indexed Date:  
2023-07-22  
Institution Address:  
Zhengzhou Univ, Sch Elect & Informat Engn, Zhengzhou 450001, Henan, Peoples R China  
Institution Address:  
Edinburgh Napier Univ, Sch Comp Engn & Built Environm, Edinburgh EH10 5DT, Scotland  
Institution:  
Institution:

### 87. A multiscale feature fusion network based on attention mechanism for

摘要: The decoding of motor imagery (MI) electroencephalogram (EEG) is an essential component of the brain- computer interface (BCI), which can help patients with motor impairment communicate directly with the outside world through assistive devices. The key to motor imagery electroencephalogram (MI-EEG) classi-fication is to extract multiple temporal, spatial, and spectral features, to obtain more comprehensive and representative information. However, current deep learning methods must fully consider the depth of temporal features and multi-spectral knowledge in EEG and often ignore the temporal or spectrum dependence in MI -EEG. In addition, the lack of effective feature fusion methods can lead to information redundancy, which affects decoding performance. To solve the above problems, this paper proposes a novel MI-EEG decoding method, named multi-scale feature fusion network based on attention mechanism (MSFF-SENet). Firstly, the multi-scale spatio-temporal module (MS-STM) and the multi-scale temporal module (MSTM) extract spatial and high-dimensional temporal features from the original signal. Then, the power spectral density estimation (PSD) convolution module (PSD-Conv module) acquires the multi-spectral features of the MI-EEG signal. Secondly, the feature fusion module fuses spatio-temporal and multi-spectral features to generate integrated feature mappings and establish the dependencies between different features. Finally, we conducted a visual analysis of the results, explaining the neural activity patterns of various motor imagery tasks in different frequency ranges and revealing the potential relationship between body movement and changes related to brain activity. The experimental results show that the classification accuracy of this model in BCI Competition IV 2a (BCI IV 2a) and High Gamma (HGD) datasets is 85.37% and 96.60%, respectively, which is superior to the most advanced methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
12  
被引频次合计:  
12  
ISSN:  
1568-4946  
eISSN:  
1872-9681  
入藏号:  
WOS:001135389400001  
Indexed Date:  
2024-01-17  
Institution Address:  
Chengdu Univ Informat Technol, Sch Comp Sci, Chengdu 610225, Peoples R China  
Institution Address:  
Univ Elect Sci & Technol China, Sch Life Sci & Technol, Chengdu 611731, Peoples R China  
Institution:  
Institution:

### 88. XctNet: Reconstruction network of volumetric images from a single X-ray

摘要: Conventional Computed Tomography (CT) produces volumetric images by computing inverse Radon trans-formation using X-ray projections from different angles, which results in high dose radiation, long reconstruction time and artifacts. Biologically, prior knowledge or experience can be utilized to identify volumetric information from 2D images to certain extents. a deep learning network, XctNet, is proposed to gain this prior knowledge from 2D pixels and produce volumetric data. In the proposed framework, self-attention mechanism is used for feature adaptive optimization; multiscale feature fusion is used to further improve the reconstruction accuracy; a 3D branch generation module is proposed to generate the details of different generation fields. Comparisons are made with the state-of-arts methods using public dataset and XctNet shows significantly higher image quality as well as better accuracy (SSIM and PSNR values of XctNet are 0.8681 and 29.2823 respectively).  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
15  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
4  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
16  
ISSN:  
0895-6111  
eISSN:  
1879-0771  
入藏号:  
WOS:000796889000002  
PubMed ID:  
35468534  
Indexed Date:  
2022-06-03  
Institution Address:  
Chinese Acad Sci, Shenzhen Inst Adv Technol, Shenzhen Key Lab Minimally Invas Surg Robot & Syst, Shenzhen 518055, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, CAS, Beijing 100049, Peoples R China  
Institution Address:  
Shenzhen Univ Clin Med Acad, Shenzhen Univ Gen Hosp, Dept Orthopaed, Shenzhen 518055, Peoples R China  
Institution:  
Institution:  
Institution:  
Shenzhen Univ Clin Med Acad

### 89. MPSU-Net: Quantitative interpretation algorithm for road cracks based on

摘要: Road cracks pose a persistent challenge in road maintenance, with timely detection and repair crucial for enhancing road safety. However, determining which cracks require repair can be difficult, necessitating a quantitative analysis approach. This paper proposes a deep learning-based method, Multiscale Feature Fusion and Superimposed U -Net (MPSU-Net), for precisely this purpose. The method employs MPSU segmentation, which quantifies crack interpretation by analyzing the black and white pixels in binary images. Within the segmentation algorithm, Attention Connection is introduced to fuse features across different layers, while the PSU Block amalgamates feature information within the same layer, incorporating ASPP, CBAM, and the superimposed U -Net. The superimposed U -Net is designed to enhance PSU's feature extraction capabilities. Furthermore, a new Conv block is introduced to bolster feature extraction by replacing all convolutions in the superimposed U -Net and decoder. To address imbalanced positive and negative crack samples, we adopt the Jaccard loss metric based on experimental results. We enhance dataset diversity by leveraging data augmentation and amalgamating data from multiple datasets, resulting in the comprehensive Crack500 and Crack datasets. Experimental findings showcase the significant efficacy of MPSU-Net in enhancing F1 -score and Recall metrics. On the Crack500 dataset, MPSU-Net achieves F1 -scores and Recall rates of 85.32% and 78.68%, respectively, representing notable improvements of 6.19% and 6.43% over the U -Net baseline performance. Similarly, on the Crack dataset, MPSU-Net attains F1 -scores and Recall rates of 74.54% and 46.66%, marking enhancements of 8.25% and 8.64% over the U -Net baseline. These results underscore the high performance of MPSU-Net and its potential to aid in quantitative road crack analysis.  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1051-2004  
eISSN:  
1095-4333  
入藏号:  
WOS:001252220800001  
Indexed Date:  
2024-06-29  
Institution Address:  
Hangzhou Dianzi Univ, Sch Mech Engn, Hangzhou 310018, Peoples R China  
Institution Address:  
Zhejiang Univ, Coll Elect Engn, Hangzhou 310027, Peoples R China  
Institution Address:  
Hangzhou City Univ, Dept Mech Engn, Hangzhou 310015, Peoples R China  
Institution:  
Institution:  
Institution:

### 90. MFCP-Net: Multiscale Feature Correlation Perception Network for Image

摘要: Due to their exceptional capacity to extract features, methods based on deep learning have demonstrated good performance in industrial vision inspection applications. However, the generalizability and robustness of inspection models are challenged by factors such as insufficient data, great differences in data types, and complex illumination. To address these problems, multiscale feature correlation perception network (MFCP-Net), a semantic segmentation model with variable topology, is proposed to segment images with prior background. MFCP-Net uses a Siamese structure with shared weights to extract the multiscale features of the template and test images. A feature correlation perception (FCP) block is designed to measure the correlation of latent representation between the template and test images at multiscales, using background information more effectively. In addition, MFCP-Net provides different topological structures for the training and testing phases. A background auto-correlation branch is included during the testing phase to eliminate false positives observed in the background. MFCP-Net was compared with the state-of-the-art models on the proposed automobile instrument detection system, demonstrating that it offers superior accuracy and generalizability, especially in the case of insufficient data and big variances across data.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000915866600014  
Indexed Date:  
2023-11-15  
Institution Address:  
Harbin Inst Technol, Res Inst Intelligent Control & Syst, Harbin 150001, Peoples R China  
Institution Address:  
Ningbo Inst Intelligent Equipment Technol Co Ltd, Ningbo 315200, Peoples R China  
Institution:  
Institution:  
Ningbo Inst Intelligent Equipment Technol Co Ltd

### 91. Automatic Crop Pest Detection Oriented Multiscale Feature Fusion

摘要: Simple Summary Monitoring pests is a labor-intensive and time-consuming task for agricultural experts. This paper proposes a new approach to classifying and counting different categories of crop pests. Specifically, we propose a multi-category pest detection network (MCPD-net), which includes a multiscale feature pyramid network and a novel adaptive feature region proposal network. The multiscale feature pyramid network is used to fuse the multiscale pest information, which significantly improves detection accuracy. The adaptive feature region proposal network addresses the problem of not aligning when region proposal network (RPN) iterating, especially for small pest objects. Extensive experiments on the multi-category pests dataset 2021 (MPD2021) demonstrated that the proposed method provides significant improvements in terms of average precision (AP) and average recall (AR); it outperformed other deep learning-based models. Specialized pest control for agriculture is a high-priority agricultural issue. There are multiple categories of tiny pests, which pose significant challenges to monitoring. Previous work mainly relied on manual monitoring of pests, which was labor-intensive and time-consuming. Recently, deep-learning-based pest detection methods have achieved remarkable improvements and can be used for automatic pest monitoring. However, there are two main obstacles in the task of pest detection. (1) Small pests often go undetected because much information is lost during the network training process. (2) The highly similar physical appearances of some categories of pests make it difficult to distinguish the specific categories for networks. To alleviate the above problems, we proposed the multi-category pest detection network (MCPD-net), which includes a multiscale feature pyramid network (MFPN) and a novel adaptive feature region proposal network (AFRPN). MFPN can fuse the pest information in multiscale features, which significantly improves detection accuracy. AFRPN solves the problem of anchor and feature misalignment during RPN iterating, especially for small pest objects. In extensive experiments on the multi-category pests dataset 2021 (MPD2021), the proposed method achieved 67.3% mean average precision (mAP) and 89.3% average recall (AR), outperforming other deep learning-based models.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
19  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
被引频次合计:  
19  
eISSN:  
2075-4450  
入藏号:  
WOS:000817373600001  
PubMed ID:  
35735891  
Indexed Date:  
2022-07-06  
Institution Address:  
Chinese Acad Sci, Hefei Inst Phys Sci, Inst Intelligent Machines, Hefei 230031, Peoples R China  
Institution Address:  
Univ Sci & Technol China, Sci Isl Branch, Grad Sch, Hefei 230026, Peoples R China  
Institution Address:  
Anhui Unviers, Sch Internet, Hefei 230039, Anhui, Peoples R China  
Institution:  
Institution:  
Institution:  
Anhui Unviers

### 92. Optical Spectral Physics-Informed Attention Network for Condition

摘要: Online condition monitoring of wire arc additive manufacturing using plasma arc spectra is essential. Purely data-driven monitoring models often cannot provide reliable classifications for out-of-sample scenarios. This article proposes an optical spectral physics-informed global-local network (GLNet) based on attention and optical spectral domain knowledge. Initially, an analysis of the physical characteristics of optical spectral signals is undertaken, with domain knowledge subsequently incorporated into the channel information during the network input construction. The architecture of GLNet includes a local branch (L-branch) to extract local structural features and a global branch (G-Branch) to consider how local patches interact globally. Additionally, this article devises a multiscale feature fusion block, which integrates multilevel features to enhance the discrimination ability of the fault condition monitoring model. In essence, the proposed network combines the power of deep learning with the generalizability of physics-based optical spectral features, enabling condition monitoring under varying conditions. Finally, the proposed method is experimentally verified on the copper and chrome zirconium copper dataset, which has considerable improvements in terms of classification accuracy.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
3  
ISSN:  
0278-0046  
eISSN:  
1557-9948  
入藏号:  
WOS:001165597300001  
Indexed Date:  
2024-03-02  
Institution Address:  
Xi An Jiao Tong Univ, Sch Elect Engn, State Key Lab Elect Insulat & Power Equipment, Xian 710049, Peoples R China  
Institution:

### 93. LW-YOLO: Lightweight Deep Learning Model for Fast and Precise Defect

摘要: Printed circuit board (PCB) manufacturing processes are becoming increasingly complex, where even minor defects can impair product performance and yield rates. Precisely identifying PCB defects is critical but remains challenging. Traditional PCB defect detection methods, such as visual inspection and automated technologies, have limitations. While defects can be readily identified based on symmetry, the operational aspect proves to be quite challenging. Deep learning has shown promise in defect detection; however, current deep learning models for PCB defect detection still face issues like large model size, slow detection speed, and suboptimal accuracy. This paper proposes a lightweight YOLOv8 (You Only Look Once version 8)-based model called LW-YOLO (Lightweight You Only Look Once) to address these limitations. Specifically, LW-YOLO incorporates a bidirectional feature pyramid network for multiscale feature fusion, a Partial Convolution module to reduce redundant calculations, and a Minimum Point Distance Intersection over Union loss function to simplify optimization and improve accuracy. Based on the experimental data, LW-YOLO achieved an mAP0.5 of 96.4%, which is 2.2 percentage points higher than YOLOv8; the precision reached 97.1%, surpassing YOLOv8 by 1.7 percentage points; and at the same time, LW-YOLO achieved an FPS of 141.5. The proposed strategies effectively enhance efficiency and accuracy for deep-learning-based PCB defect detection.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
eISSN:  
2073-8994  
入藏号:  
WOS:001210712000001  
Indexed Date:  
2024-06-29  
Institution Address:  
East China Jiaotong Univ, Sch Software, Nanchang 330013, Peoples R China  
Institution:

### 94. DMFR-YOLO: an infrared small hotspot detection algorithm based on double

摘要: Hotspot automatic detection is an effective strategy to realize intelligent maintenance of photovoltaic (PV) modules. However, it is challenging to detect small hotspots accurately using deep learning-based object detection methods due to the weak information and small area of the target. This study presents a double multi-scale feature reconstruction (DMFR)-YOLO to detect small hotspots in PV modules. In the proposed DMFR-YOLO, the backbone of YOLOv8n is improved by eliminating redundant deep layers and introducing a shallow detection head to enhance its perception and recognition ability for small objects. Moreover, a weighted multi-layer feature reconstruction (MLFR) module is introduced to fuse the features of different depths and a multi-receptive field feature reconstruction (MRFR) module is designed to fuse the information of different receptive fields. Finally, ablation and comparison experiments are conducted to evaluate the performance of the proposed method. The model with both MLFR and MRFR can achieve a recall rate of 88.6%, with a mean average precision (mAP@0.5) of 93.3%. Experimental results demonstrate that the proposed DMFR-YOLO is able to achieve small hotspot detection in IR images with high accuracy.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001358781600001  
Indexed Date:  
2024-11-26  
Institution Address:  
North China Elect Power Univ, Sch Control & Comp Engn, Beijing, Peoples R China  
Institution:

### 95. SNLRUX plus plus for Building Extraction From High-Resolution Remote

摘要: Building extraction plays an important role in high-resolution remote sensing image processing, which can be used as the basis for urban planning and demographic analysis. In recent years, many powerful general semantic segmentation models have emerged, but these models often perform poorly when transferred to remote sensing images because of the characteristics of remote sensing images. To this end, we propose a new deep learning network called Selective Nonlocal ResUNeXt++ (SNLRUX++) for building extraction. First, the cascaded multiscale feature fusion is proposed to transform the high-performance image classification network ResNeXt into the segmentation network ResUNeXt++. Second, selective nonlocal operation is designed to establish long-range dependencies while avoiding introducing excessive noise and computational effort. Finally, multiscale prediction is applied as deep supervision to accelerate training and convergence, and improves prediction performance of objects at different scales. The experimental results on two different remote sensing image datasets show the effectiveness and generalization ability of the proposed method.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
14  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000736740000006  
Indexed Date:  
2023-05-23  
Institution Address:  
Zhejiang Univ Technol, Coll Comp Sci & Technol, Hangzhou 310023, Peoples R China  
Institution:

### 96. A Novel Method for Segmentation and Detection of Weld Defects in UHV

摘要: A novel method for detecting weld defects in ultra-high voltage (UHV) equipment is present by combining unimodal semantic segmentation with X-ray imaging. The approach begins by employing a deep neural network to extract weak weld features from X-ray images. A channel attention module is introduced to balance the importance of different feature weights, enhancing the network's ability to focus on key features. An atrous spatial pyramid pooling module is then utilized to expand the receptive field, effectively leveraging the spatial hierarchical information within the X-ray images. Additionally, a multi-scale feature fusion module is applied to automatically learn feature relationships, capturing semantic information at various scales, which significantly improves the distinction between defective and normal weld regions. The method's effectiveness is validated through repeated experiments on the GDXray weld dataset and a self-constructed UHV weld dataset. Quantitative comparisons demonstrate that the proposed method significantly enhances the segmentation accuracy of weld defects in UHV equipment, providing a valuable tool for technicians in the field of weld non-destructive testing.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1061-8309  
eISSN:  
1608-3385  
入藏号:  
WOS:001407150800009  
Indexed Date:  
2025-02-03  
Institution Address:  
China Univ Min & Technol, Sch Mech & Elect Engn, Jiangsu Key Lab Mine Mech & Elect Equipment, Xuzhou 221116, Jiangsu, Peoples R China  
Institution Address:  
China Railway Guangzhou Bur Grp Co Ltd, Changsha Railway Logist Ctr, Changsha 410000, Hunan, Peoples R China  
Institution Address:  
Henan Pinggao Elect Co Ltd, China Elect Equipment Grp Co Ltd, Pingdingshan 467001, Henan, Peoples R China  
Institution Address:  
UCL, Dept Phys & Astron, Gower St, London WC1E 6BT, England  
Institution Address:  
Zhengzhou Univ Engn & Technol, Sch Informat Engn, Zhengzhou 450044, Henan, Peoples R China  
Institution:  
Institution:  
China Railway Guangzhou Bur Grp Co Ltd  
Institution:  
Henan Pinggao Elect Co Ltd  
Institution:  
Institution:  
Zhengzhou Univ Engn & Technol

### 97. Feedback Multi-scale Residual Dense Network for image super-resolution

摘要: The image super-resolution algorithm based on deep learning has a good reconstruction effect, and the reconstruction can be further enhanced by using multi-scale features. There are different extraction methods for multi-scale features, and current deep learning-based super-resolution algorithms often use only one method when utilizing multi-scale features. We use an error feedback mechanism with a dense residual mechanism to fuse multi-scale features and propose Feedback Multi-scale Residual Dense Network (FMDN), which uses two different multi-scale features to enhance the reconstruction effect. On the other hand, in the previous multi-scale feature fusion often used the method of concatenating. We design a new error feedback-based feature fusion method, and the experimental results show that it has better results than the common method of concatenating. In addition, we further use the feedback mechanism of recurrent to improve the efficiency of the module, which can use fewer layers to achieve the effect of more layers of the basic model, and take up less space, faster, or make a network with a larger number of layers have better results. Compared with the state-of-the-art method, the proposed method shows promising performance according to qualitative and quantitative evaluation.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0923-5965  
eISSN:  
1879-2677  
入藏号:  
WOS:000823352000008  
Indexed Date:  
2022-07-20  
Institution Address:  
Guangdong Polytech Normal Univ, Sch Comp Sci, 293,Zhongshan Ave West,Tianhe Dist, Guangzhou 510665, Guangdong, Peoples R China  
Institution:

### 98. BFEA: A SAR Ship Detection Model Based on Attention Mechanism and

摘要: With the advancements in deep learning and synthetic aperture radar (SAR) technology, an increasing number of individuals are utilizing deep-learning techniques to detect ships in SAR images. However, the efficiency of SAR ship detection is affected by complex background interference and various ship sizes. Addressing these challenges, this article proposes a balanced feature enhanced attention model. First, we introduce a novel attention feature fusion network (WEF-Net) tailored for SAR multiscale ship detection. WEF-Net effectively balances the information across different backbone layers and harmonizes semantic information from various levels of the feature pyramid through aggregation and averaging. Next, we embed the receiving field extension module in WEF-Net to learn the context information and generate the global characteristics of the receiving field balance. In addition, it can extract features from multiple scales to enhance the detection capability of the model for ships of different scales. At the same time, acknowledging the impact of surrounding complex background interference on the detector, we redesigned the ELAN module by combining convolution and attention. This enhancement enables the model to better attend to target position information during feature fusion, suppress the surrounding complex background interference, and highlight the ship's feature information. Finally, owing to the prevalence of small targets in SAR images, we employ an optimized loss function to bolster the model's performance in detecting small targets. This approach accelerates training convergence, reduces instances of missed detection on small targets, and enhances overall detection performance across multiple scales. Experimental results demonstrate that our model achieves detection accuracies of 98%, 93.1%, and 76.9% on the SAR ship detection dataset, high-resolution SAR image dataset, and large-scale SAR ship detection dataset, respectively, effectively discerning ship targets amid complex backgrounds in SAR images.  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001252509000008  
Indexed Date:  
2024-07-06  
Institution Address:  
Henan Univ, Key Lab Big Data Anal & Proc, Kaifeng 475000, Peoples R China  
Institution Address:  
Henan Univ, Sch Comp & Informat Engn, Kaifeng 475000, Peoples R China  
Institution:  
Institution:

### 99. MLFNet- Point Cloud Semantic Segmentation Convolution Network Based on

摘要: In the semantic segmentation of a point cloud, if the spatial structure correlation between the input features and coordinates are not fully considered, a semantic segmentation error can occur. We propose a method of spatial convolution that makes full use of the characteristics of a multiscale spatial structure by combining local and global features. We call this method MLFNet. We also propose a multiscale feature framework. First, the point cloud is simplified by obtaining the weighted farthest point (by down-sampling combined with farthest-point sampling and the weighted average). The near-near domain of each sampling point is then obtained by a KK octant search (an octant search optimized by the k-nearest neighbor and a custom threshold), and feature information is obtained. The feature information is added to the subsequent multilayer perceptron, and fusion of local context information is achieved. Finally, the fusion features in multiple directions are maximally pooled. Our method was tested on self-made datasets and other standard basic datasets (ModelNet40, ShapeNet, and Stanford large-scale 3D indoor spaces (S3DIS) data). The accuracy of segmentation was 0.937 in our dataset; two percentage points higher than the latest deep learning method. Also, our method obtained a mean intersection over a union of 0.867 in ShapeNet, which was 0.3 percentage points higher than the latest PointGrid. The accuracy on S3DIS was 0.8153, which was three percentage points higher than the latest spatial aggregation net. The results of semantic segmentation verified the superiority of the proposed method.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
3  
ISSN:  
2169-3536  
入藏号:  
WOS:000633371700001  
Indexed Date:  
2021-04-01  
Institution Address:  
Qilu Univ Technol, Sch Mech & Automot Engn, Shandong Acad Sci, Jinan 250300, Peoples R China  
Institution:

### 100. A novel difficult-to-segment samples focusing network for oral CBCT

摘要: Using deep learning technology to segment oral CBCT images for clinical diagnosis and treatment is one of the important research directions in the field of clinical dentistry. However, the blurred contour and the scale difference limit the segmentation accuracy of the crown edge and the root part of the current methods, making these regions become difficult-to-segment samples in the oral CBCT segmentation task. Aiming at the above problems, this work proposed a Difficult-to-Segment Focus Network (DSFNet) for segmenting oral CBCT images. The network utilizes a Feature Capturing Module (FCM) to efficiently capture local and long-range features, enhancing the feature extraction performance. Additionally, a Multi-Scale Feature Fusion Module (MFFM) is employed to merge multiscale feature information. To further improve the loss ratio for difficult-to-segment samples, a hybrid loss function is proposed, combining Focal Loss and Dice Loss. By utilizing the hybrid loss function, DSFNet achieves 91.85% Dice Similarity Coefficient (DSC) and 0.216 mm Average Symmetric Surface Distance (ASSD) performance in oral CBCT segmentation tasks. Experimental results show that the proposed method is superior to current dental CBCT image segmentation techniques and has real-world applicability.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2045-2322  
入藏号:  
WOS:001179043600022  
PubMed ID:  
38429362  
Indexed Date:  
2024-03-31  
Institution Address:  
Zhejiang Shuren Univ, Coll Informat Sci & Technol, Hangzhou 310015, Peoples R China  
Institution Address:  
Zhejiang Shuren Univ, Zhejiang Netherlands Joint Lab Digital Diag & Trea, Hangzhou 310015, Zhejiang, Peoples R China  
Institution:  
Institution:

### 101. EMAFF-Net: an enhanced multi-scale attentive feature fusion network for

摘要: Automated building extraction is imperative for several geospatial applications such as monitoring disaster-affected buildings and urban planning. Existing deep learning (DL)-based building extraction methods fail to capture high-level semantic features due to the complex nature and diverse appearance of visually similar structures. To address this issue, in this letter, we propose an enhanced multi-scale attentive feature fusion network (EMAFF-Net) for building extraction from remote sensing (RS) images. EMAFF-Net is an end-to-end DL architecture based on U-Net that includes: i) an encoder; ii) an enhanced multi-scale feature fusion (EMFF) module; iii) a refined multi-scale convolutional block attention (RM-CBAM) module and iv) a decoder with refinement layers. To extract multi-scale contextual information, we incorporate an RM-CBAM module into the lateral connections of encoder-decoder layers of EMAFF-Net. Further, a novel EMFF module is integrated to obtain fine-grained features from the lowest encoder layer with minimal trainable parameters required. We evaluate the performance of the proposed network on two benchmark datasets: Massachusetts (MAS) and WHU building datasets. The experimental results show that the proposed approach outperforms the existing reference methods showcasing its potential in practical applications.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
2150-704X  
eISSN:  
2150-7058  
入藏号:  
WOS:001151002000001  
Indexed Date:  
2024-02-05  
Institution Address:  
Amrita Vishwa Vidyapeetham, Amrita Sch Comp, Dept Comp Sci & Engn, Kollam, India  
Institution:

### 102. Multi-scale feature fusion kernel estimation with masked interpolation

摘要: In recent years, the application of deep learning to remote sensing image super-resolution has achieved promising results. However, most deep learning-based methods are often trained on remote sensing datasets constructed by bicubic downsampling, and their recovery effects on real-world remote sensing images are often unsatisfactory. This is because the process of generating low-resolution (LR) images from high-resolution (HR) images to construct training data pairs (LR-HR) using simple bicubic downsampling cannot reflect the degradation process of real-world remote sensing dataset images. In this paper, we propose a multi-scale feature fusion kernel estimation with masked interpolation loss for real-world remote sensing images super-resolution (MFFILSR) to address this problem. MFFILSR is divided into two stages: degenerate kernel estimation and SR network training. In the first stage, we propose a multi-scale feature fusion kernel estimation network that can effectively fuse multi-scale information, making the estimated downsampling kernel closer to the degradation patterns of real-world remote sensing images. In the second stage, we introduce a masked interpolation loss during generator training, by masking the interpolation loss, the artefacts of generated images can be effectively reduced. Extensive experiments show that MFFILSR has satisfactory super-resolution reconstruction performance for real-world remote sensing images.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:001064402900001  
Indexed Date:  
2023-09-23  
Institution Address:  
China Univ Petr East China, Coll Control Sci & Engn, Qingdao 266580, Shandong, Peoples R China  
Institution Address:  
China Univ Petr East China, Coll Oceanog & Space Informat, Qingdao, Shandong, Peoples R China  
Institution Address:  
Univ Macau, Fac Sci & Technol, Dept Comp & Informat Sci, Fg, Macau, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Acad Optoelect, Beijing, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 103. RSI-YOLO: Object Detection Method for Remote Sensing Images Based on

摘要: With the continuous development of deep learning technology, object detection has received extensive attention across various computer fields as a fundamental task of computational vision. Effective detection of objects in remote sensing images is a key challenge, owing to their small size and low resolution. In this study, a remote sensing image detection (RSI-YOLO) approach based on the YOLOv5 target detection algorithm is proposed, which has been proven to be one of the most representative and effective algorithms for this task. The channel attention and spatial attention mechanisms are used to strengthen the features fused by the neural network. The multi-scale feature fusion structure of the original network based on a PANet structure is improved to a weighted bidirectional feature pyramid structure to achieve more efficient and richer feature fusion. In addition, a small object detection layer is added, and the loss function is modified to optimise the network model. The experimental results from four remote sensing image datasets, such as DOTA and NWPU-VHR 10, indicate that RSI-YOLO outperforms the original YOLO in terms of detection performance. The proposed RSI-YOLO algorithm demonstrated superior detection performance compared to other classical object detection algorithms, thus validating the effectiveness of the improvements introduced into the YOLOv5 algorithm.  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
23  
被引频次合计:  
23  
eISSN:  
1424-8220  
入藏号:  
WOS:001038903600001  
PubMed ID:  
37514708  
Indexed Date:  
2023-08-10  
Institution Address:  
Northeast Elect Power Univ, Sch Comp Sci, Jilin 132012, Peoples R China  
Institution Address:  
Northeast Elect Power Univ, Sch Energy & Power Engn, Jilin 132012, Peoples R China  
Institution:  
Institution:

### 104. Camouflaged Object Detection via location-awareness and feature fusion

摘要: Camouflaged object detection aims to completely segment objects immersed in their surroundings from the background. However, existing deep learning methods often suffer from the following shortcomings: (1) They have difficulty inaccurately perceiving the target location; (2) The extraction of multi-scale feature is insufficient. To address the above problems, we proposed a camouflaged object detection network(LFNet) based on location-awareness and feature fusion. Specifically, we designed a status location module(SLM) that dynamically captures the structural features of targets across spatial and channel dimensions to achieve accurate segmentation. Beyond that, a residual feature fusion module(RFFM) was devised to address the challenge of insufficient multi-scale feature integration. Experiments conducted on three standard datasets(CAMO,COD10K and NC4K) demonstrate that LFNet achieves significant improvements compared with 15 state-of-the-art methods. The code will be available at https://github.com/ZX123445/LFNet.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0262-8856  
eISSN:  
1872-8138  
入藏号:  
WOS:001361988100001  
Indexed Date:  
2024-12-03  
Institution Address:  
Northeast Petr Univ, Dept Elect Informat Engn, Daqing 163318, Peoples R China  
Institution Address:  
Daqing Branch China Mobile Commun Grp Heilongjiang, Daqing 163318, Peoples R China  
Institution Address:  
China Univ Petr East China, Sch Comp Sci & Technol, Qingdao 266580, Peoples R China  
Institution:  
Institution:  
Daqing Branch China Mobile Commun Grp Heilongjiang  
Institution:

### 105. MAL-YOLO: a lightweight algorithm for target detection in side-scan

摘要: Side-scan sonar image target detection is of great significance in seabed resource exploration and other fields. However, affected by the complex underwater environment, side-scan sonar images have the problems of few target samples and large differences in the scale of each type of target. In addition, the computational complexity of high-performance models based on deep learning is too high to be applied on platforms with limited computational resources. To solve these problems, this paper proposes a lightweight algorithm for target detection in side-scan sonar images based on multi-scale feature fusion and attention mechanism (MAL-YOLO). Firstly, a lightweight feature extraction module is used. This module combines depthwise separable convolution and efficient multi-scale attention (EMA) module to improve the feature extraction capability of the model while reducing the computational volume. Secondly, a multi-scale feature fusion network combining asymptotic feature pyramid network and EMA module is used to enhance the fusion and representation of multi-scale features in the model. Finally, the MPDIoU loss function is used to provide more accurate bounding box regression. The experimental results show that the algorithm has significant advantages in both detection accuracy and model lightweighting compared with the current state-of-the-art algorithms such as YOLOv7 and YOLOv8.  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1753-8947  
eISSN:  
1753-8955  
入藏号:  
WOS:001304464000001  
Indexed Date:  
2024-09-08  
Institution Address:  
Shandong Univ Sci & Technol, Coll Geodesy & Geomat, Qingdao 266590, Shandong, Peoples R China  
Institution Address:  
Minist Nat Resources, Key Lab Surveying & Mapping Technol Isl & Reef, Qingdao 266590, Shandong, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Acoust, Beijing 100190, Peoples R China  
Institution Address:  
Unit 91001, Beijing, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Unit 91001

### 106. Camouflage soldier object detection network based on the attention

摘要: Due to the high level of information similarity between camouflage soldier objects and their background, traditional deep learning-based object detection networks encounter distinct error detection rates and miss detection rates when attempting to detect camouflage soldiers. To address these challenges, we proposed a camouflage soldier object detection network (AFSNet) based on attention mechanism and multi-scale feature fusion strategy. We employed an attention module to enhance the network's capability for feature extraction. Furthermore, we proposed a novel strategy for multi-scale feature fusion based on pyramidal feature shrinking, aiming to mitigate interference caused by interpolation and prevent information loss resulting from pooling during the process of feature fusion. Moreover, we introduced a novel information handle module that enhances the network's capability for feature fusion by regulating the information transmission pathway. Experiments demonstrated that our network exhibits a better camouflage object detection performance than state-of-arts networks. Compared to YOLOv7, our network can achieve 93%\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\%$$\end{document} AP, which is increased by 6.7%\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\%$$\end{document} with almost no computation overhead.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001172078800009  
Indexed Date:  
2024-04-25  
Institution Address:  
Beijing Inst Technol, Sch Mechatron Engn, Beijing 100081, Peoples R China  
Institution:

### 107. Freshness uniformity measurement network based on multi-layer feature

摘要: The arrangement of products on supermarket freshness shelves exhibits a certain pattern and displays distinct texture characteristics. In recent years, many studies have applied texture extraction algorithms in deep learning, such as the Histogram Layer Residual Network (HistNet). However, this algorithm still has obvious disadvantages, such as neglecting the optimal representation of multi-scale texture features and lacking feature selection during extraction. To address these issues, this paper introduces a novel texture classification network-Multi-Scale Feature Histogram Network (MFHisNet). First, we design a Multi-Scale Feature FusionModule (MF-Block) to achieve a multi-level representation of texture information. Then, we utilize an attention module (CBAM) to weight crucial information and suppress background interference for deeper level texture features. Experimental results demonstrate that the model achieves accuracies of 82.12 +/- 2.04%, 73.13 +/- 1.10%, and 83.46 +/- 0.62% on theGTOS-mobile, DTD, and MINC-2500 datasets, respectively. Furthermore, based on the proposed model, we propose a measurement method that uses cosine similarity to measure the uniformity of freshness placement, and the effectiveness of this method was verified on the dataset we collected.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001119988100002  
Indexed Date:  
2023-12-23  
Institution Address:  
Huzhou Univ, Sch Informat Engn, Huzhou 313000, Peoples R China  
Institution Address:  
Zhejiang Univ, Huzhou Inst, Huzhou 313000, Peoples R China  
Institution Address:  
Liaoning Normal Univ, Sch Comp & Informat Technol, Dalian 116081, Peoples R China  
Institution:  
Institution:  
Institution:

### 108. HRLE-SARDet: A Lightweight SAR Target Detection Algorithm Based on

摘要: In recent years, deep learning has been widely used in remote sensing, especially in the field of synthetic aperture radar (SAR) image target detection. However, all of these deep learning models continue increasing the network's depth and width without maintaining a good balance between accuracy and speed. Therefore, in this article, we propose a hybrid representation learning-enhanced SAR target detection algorithm based on the unique features of SAR images from a lightweight perspective called HRLE-SARDet. First, we design a lightweight and scattering feature extraction backbone that is more suitable for SAR image data. Second, for the multiscale feature discrepancy, we design a new multiscale feature fusion neck. Next, to better extract the scattering information from small targets of SAR images and improve the detection accuracy, we design a lightweight hybrid representation learning enhancement module. Finally, to better fit target detection for SAR image datasets, we redesign a more flexible loss function, which allows for an easy adjustment of the importance of polynomial bases according to the target task and dataset. Extensive experimental results on three SAR image ship target datasets (SSDD, AIR-SARShip-2.0, and HRSID) and a newly released large multiclass target SAR dataset (MSAR-1.0) show that our HRLE-SARDet achieves 98.4%, 79.2%, 92.5%, and 88.4% mean average precision (mAP) with only 1.09 M parameters and 2.5 G floating-point operations (FLOPs) on the SSDD, AIR-SARShip-2.0, HRSID, and MSAR-1.0 datasets, respectively, which is an excellent performance.  
Web of Science 核心合集中的 "被引频次":  
39  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
39  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000951955500009  
Indexed Date:  
2023-04-06  
Institution Address:  
Anhui Univ, Sch Elect & Informat Engn, Informat Mat & Intelligent Sensing Lab Anhui Prov, Key Lab Intelligent Comp & Signal Proc Minist Educ, Hefei 230601, Peoples R China  
Institution Address:  
China Elect Technol Grp Corp, Res Inst 38, Hefei 230088, Peoples R China  
Institution Address:  
Univ Fed Rio Janeiro, COPPE Poli, Program Elect Engn, Dept Elect & Comp Engn, BR-21941972 Rio De Janeiro, RJ, Brazil  
Institution:  
Institution:  
Institution:  
Univ Fed Rio Janeiro

### 109. Mural inpainting with generative adversarial networks based on

摘要: This study proposes a deep learning model for mural restoration based on generative adversarial networks with multi-scale feature and attention fusions, addressing insufficient feature extraction and detail loss of the existing deep learning image inpainting algorithms during reconstruction. Firstly, a multi-scale feature pyramid network is designed to extract feature information of different scales in mural images, which enhances the feature relevance. Secondly, using the self-attention mechanism and feature fusion module, a multi-scale feature generator is constructed to obtain rich context information and improve the restoration ability of the network. Finally, the minimal confrontation loss and the mean square error are introduced to promote the residual feedback of the discriminator, which completes the mural restoration by combining the feature information of different scales. The experimental results of digital restoration of real Dunhuang murals show that the proposed algorithm can effectively protect important feature information such as the edges and textures, and that the subjective visual effects and objective evaluation indicators are superior to those of the algorithms for comparison.  
摘要:  
针对现有深度学习图像修复算法修复壁画时,存在特征提取不足及细节重构丢失等问题,提出了一种多尺度特征和注意力融合的生成对抗壁画修复深度学习模型。设计多尺度特征金字塔网络提取壁画中不同尺度的特征信息,增强特征关联性;采用自注意力机制及特征融合模块构建多尺度特征生成器,以获取丰富的上下文信息,提升网络的修复能力;引入最小化对抗损失与均方误差促进判别器的残差反馈,从而结合不同尺度的特征信息完成壁画修复。通过对真实敦煌壁画数字化修复的实验结果表明,所提算法能够有效保护壁画图像的边缘和纹理等重要特征信息,并且主观视觉效果及客观评价指标均优于比较算法。  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1001-5965  
入藏号:  
CSCD:7430278  
Indexed Date:  
2023-07-01  
Institution Address:  
兰州交通大学电子与信息工程学院;;甘肃省人工智能与图形图像处理工程研究中心, ;;甘肃省人工智能与图形图像处理工程研究中心, 兰州;;兰州, ;; 730070;;730070  
Institution Address:  
兰州交通大学电子与信息工程学院, 兰州, 甘肃 730070, 中国  
Institution Address:  
School of Electronic and Information Engineering,Lanzhou Jiaotong University;;Gansu Provincial Engineering Research Center for Artificial Intelligence and Graphics & Image Processing, ;;Gansu Provincial Engineering Research Center for Artificial Intelligence and Graphics & Image Processing, Lanzhou;;Lanzhou, ;; 730070;;730070  
Institution Address:  
School of Electronic and Information Engineering,Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China  
Institution:  
兰州交通大学电子与信息工程学院;;甘肃省人工智能与图形图像处理工程研究中心  
Institution:  
兰州交通大学电子与信息工程学院  
Institution:  
School of Electronic and Information Engineering,Lanzhou Jiaotong University;;Gansu Provincial Engineering Research Center for Artificial Intelligence and Graphics & Image Processing  
Institution:  
School of Electronic and Information Engineering,Lanzhou Jiaotong University

### 110. UW-DETR: Feature Fusion Enhanced RT-DETR for Improving Underwater Object

摘要: Deep learning models for object detection have shown favorable results in controlled environments. However, owing to the complexity of underwater environments, these models suffer from target blurring caused by noise, as well as visual disturbances caused by the protective coloration of underwater organisms. These methods are not sufficiently equipped to effectively address the challenge of underwater object detection. To address the challenges in detection caused by underwater blurriness and the camouflage of objects. In this study, a method called Spatial Semantic Encoding Fusion(SSEF) is proposed. Pooling and bilinear interpolation are utilized by SSEF to standardize features across different scales. The standardized features are then fused using the Hadamard product. The fused features are encoded separately for spatial information and channel semantic information through a dual-branch structure. Finally, the dual-branch features are combined by point-wise addition. SSEF is utilized to enhance the multi-scale feature fusion approach of RT-DETR-r18, proposing an underwater object detection framework named Underwater-DETR(UW-DETR). The experiments are conducted using the UTDAC2020 and Brackish datasets. The experimental results show that UW-DETR outperforms other underwater object detection methods and UW-DETR meets the requirements for underwater applications.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001381377200013  
Indexed Date:  
2025-01-05  
Institution Address:  
Shandong Jiaotong Univ, Naval Architecture & Port Engn Coll, Weihai 264200, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automat, Beijing 100190, Peoples R China  
Institution:  
Institution:

### 111. White blood cell classification network using MobileNetv2 with

摘要: White blood cells play a crucial role in the human immune system. The accurate classification of white blood cells can help doctors diagnose various diseases for patients. To enhance the classification accuracy of white blood cells micro-vision images, an efficient lightweight deep learning network called ICAFF-MobileNetv2 is proposed in this paper. Firstly, the pruning operations are applied to the MobileNetv2 and the multiscale feature extraction module is proposed to reduce the model size while ensuring the classification accuracy. Secondly, the improved coordinate attention mechanism module and the attention feature fusion module are incorporated into the inverted residual structures to enhance the feature extraction capability and the classification accuracy of the network. Furthermore, considering the imbalance among different types of white blood cells in the datasets, the method of label smoothing is introduced to improve the cross-entropy loss function and further enhance the classification accuracy. Finally, the improved network is validated on the proposed mixed dataset. The experimental results demonstrate that the proposed network features high classification accuracy and small model size. Compared to other classification networks, ICAFF-MobileNetv2 exhibits superior performance with 98.54 %, 98.21 %, 98.56 %, and 98.38 % of the classification accuracy, recall, precision and F1 score respectively, which are 0.7 %, 0.86 %, 0.92 %, and 0.89 % better than MobileNetv2.  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:001315479600001  
Indexed Date:  
2024-09-25  
Institution Address:  
Hunan Univ Sci & Technol, Sch Informat & Elect Engn, Xiangtan 411201, Peoples R China  
Institution:

### 112. Fault classification method of operation and maintenance knowledge base

摘要: Traditional operation and maintenance knowledge base does not have the ability to identify the failure phenomena in the image. Therefore, the knowledge base cannot handle the problem of unstructured data. To tackle this issue, based on fault classification networks in deep learning, an improved capsule network feature extraction structure based Caps-DRFN algorithm is proposed, which can realize automatic classification of operation and maintenance images of electromechanical equipment. Firstly, aiming at the multi-noise problem of operation and maintenance images, the deep residual shrinkage networks (DRSN) are introduced to improve the feature extraction performance of the model on noisy data. Subsequently, for the multi-scale problem of actual shooting operation and maintenance images, through the combination of the feature pyramid networks (FPN) algorithm, the Caps-DRFN realizes image multiscale feature fusion and improves the accuracy of model classification. Finally, the vector neuron is constructed by using the capsule structure, and the digital capsule of the classification structure is obtained through the feature transmission method of dynamic routing. The model realizes the fault classification of electromechanical equipment. The experimental results show that compared with the traditional capsule network algorithm, the accuracy of the proposed Caps-DRFN algorithm based on feature fusion is increased by 15% and it is more robust.  
摘要:  
针对传统运维知识库不具备图像故障现象识别能力,无法处理非结构化数据的问题,基于深度学习的故障分类网络,提出改进胶囊网络特征提取结构的Caps-DRFN算法,实现机电设备运维图像自动分类。首先,针对运维图像存在的多噪声问题,引入深度残差收缩网络(deep residual shrinkage networks,DRSN)提高模型在含噪声数据上的特征提取效果。然后,针对实际拍摄的运维图像多尺度问题,结合FPN(feature pyramid networks)算法,实现图像多尺度特征融合提高模型分类准确率。最后,利用胶囊结构构建向量神经元,通过动态路由的特征传递方式,得到分类结构数字胶囊,实现机电设备故障分类。实验结果表明,相较于传统胶囊网络算法,提出的基于特征融合的Caps-DRFN算法准确率提高了15%且有着更强的鲁棒性。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1000-7105  
入藏号:  
CSCD:7251737  
Indexed Date:  
2022-09-30  
Institution Address:  
江南大学物联网工程学院, 无锡, 江苏 214122, 中国  
Institution Address:  
School of Internet of Things Engineering, Jiangnan University, Wuxi, Jiangsu 214122, China  
Institution:  
江南大学物联网工程学院  
Institution:  
School of Internet of Things Engineering, Jiangnan University

### 113. Integrating sentinel-2a imagery, DEM data, and spectral feature analysis

摘要: Landslides can cause severe damage to property and human life. Identifying their locations and characteristics is crucial for emergency rescue and disaster risk assessment. However, existing methods need help in accurately detecting landslides because of their diverse characteristics and scales, as well as the differences in spectral features and spatial heterogeneity of remote sensing images. To overcome these challenges, a multiscale feature fusion landslide-detection network (MFLD-Net) is proposed. This network utilizes reflectance difference images from pre- and post-landslide Sentinel-2A images, along with digital elevation model (DEM) data. Moreover, a multichannel differential landslide dataset was constructed through spectral analysis of Sentinel-2A images, which facilitates network training and enables differentiation between landslides and other objects with similar spectral features, such as bare soil and buildings. The proposed MFLD-Net was tested in Shuzheng Valley and Detuo town in Sichuan, China, where earthquakes have occurred. The experimental results revealed that compared with advanced deep learning models, MFLD-Net has promising landslide detection performance. This study provides suggestions for selecting optimal deep learning methods and spectral band combinations for landslide detection and offers a publicly available landslide dataset for further research.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1612-510X  
eISSN:  
1612-5118  
入藏号:  
WOS:001329816800001  
Indexed Date:  
2024-10-17  
Institution Address:  
South China Normal Univ, Sch Geog Sci, Guangzhou 336099, Peoples R China  
Institution Address:  
Shandong Univ Sci & Technol, Sch Surveying Mapping & Spatial Informat, Qingdao 266590, Peoples R China  
Institution Address:  
South China Normal Univ, BeiDou Res Inst, Sch Geog Sci, Foshan 528225, Peoples R China  
Institution Address:  
Shandong Prov Inst Land Surveying & Mapping, Jinan 250100, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Shandong Prov Inst Land Surveying & Mapping

### 114. A Prior-Guided Face Image Super -Resolution Network Based on Attention

摘要: In this paper, we propose a face image super-resolution method, aiming to reconstruct high quality face images from a low resolution input. The proposed method introduces an attentional multi-scale feature fusion block, which aims to improve the representation power of the neural network by emphasizing the important feature maps and suppressing the unimportant ones. In addition, the facial prior information is utilized by adding a separate prior branch, an hourglass structure is used. Experiments show that the face images reconstructed by the proposed method exhibit noticeable quality improvement compared to the low-resolution images and other SR approaches.  
会议名称:  
International Workshop on Advanced Imaging Technology (IWAIT)  
会议日期:  
JAN 09-11, 2023  
会议地点:  
Jeju, SOUTH KOREA  
会议赞助商:  
Korean Inst Broadcast & Media Engineers; Sungkyunkwan Univ, BK21 Four Educ & Res Inst Sustainable ICT Future Human Resources; Inst Elect, Informat & Commun Engineers; Inst Image Informat & Televis Engineers; Japanese Soc Precis Engn  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0277-786X  
eISSN:  
1996-756X  
ISBN:  
978-1-5106-6308-4; 978-1-5106-6309-1  
入藏号:  
WOS:001004075700038  
Indexed Date:  
2023-07-23  
Institution Address:  
Waseda Univ, 3-14-9 Okubo,Shinjuku Ku, Tokyo, Japan  
Institution Address:  
Sharp Co Ltd, I-9-2Mihama Ku, Chiba, Japan  
Institution:  
Institution:

### 115. Study on Building Extraction from Remote Sensing Image Based on

摘要: Building extraction from remote sensing images based on deep learning has the characteristics of wide coverage and high computational efficiency,and it plays an important role in urban construction,disaster prevention and other aspects.Most of the mainstream methods use multi-scale feature fusion to enable the neural network to learn more abundant semantic information. However,due to the complexity of multi-scale features and the interference of other ground objects,this kind of methods often lead to target missing and noise-intensive.To this end,this paper proposes a feature interpretation model MGA-ResNet50 (MGAR)that combines attention mechanism.The core of the method is to use the multihead attention to process the hierarchical weighting of high-level semantic information,so as to extract the optimal feature combination with relatively better representation effect.Then use the gating structure to fuse the feature map of each dimension with the low-level semantic information of the corresponding encoder to compensate for the loss of local building details.Experimental results on public datasets such as Massachusetts Building and WHU Building show that the proposed algorithm can achieve higher F1and IoU than the more advanced multiscale feature fusion methods such as RAPNet,GAMNet and GSM.  
摘要:  
基于深度学习的遥感影像建筑物提取方法具有覆盖范围广、运算效率高的特点,在城市建设、灾害防治等方面有着重要的实际意义。主流方法大多采用多尺度特征融合的方式使神经网络能够学习到更丰富的语义信息,然而由于受到多尺度特征的复杂性以及其他类别地物的干扰,该类方法往往存在着目标漏检与噪声密集的问题。对此,文中设计并实现了一种结合注意力机制的特征解译模型MGA-ResNet50(MGAR)。该方法的核心在于利用多头注意力对高等级语义信息进行分层加权处理,以提取出表征效果较好的最优特征组合;而后使用门控结构将每维特征图与对应编码端的低级语义信息融合,来解决局部建筑物细节信息丢失的问题。在Massachusetts Building,WHU Building等公开数据集上的实验结果表明,与RAPNet,GAMNet, GSM等较为先进的多尺度特征融合方法相比,所提算法能够取得更高的F1与IoU指标。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1002-137X  
入藏号:  
CSCD:7720389  
Indexed Date:  
2024-07-26  
Institution Address:  
郑州大学地球科学与技术学院, 郑州, 河南 450052, 中国  
Institution Address:  
郑州大学计算机与人工智能学院, 郑州, 河南 450001, 中国  
Institution Address:  
School of Earth Science and Technology,Zhengzhou University, Zhengzhou, Henan 450052, China  
Institution Address:  
School of Computer and Artificial Intelligence,Zhengzhou University, Zhengzhou, Henan 450001, China  
Institution:  
郑州大学地球科学与技术学院  
Institution:  
郑州大学计算机与人工智能学院  
Institution:  
School of Earth Science and Technology,Zhengzhou University  
Institution:  
School of Computer and Artificial Intelligence,Zhengzhou University

### 116. A Multiscale Feature Fusion Method for Automatic Detection of Eggs From

摘要: The widespread Pomacea canaliculata and Pomacea maculata in North America and Asia have caused significant adverse effects on the local ecological environment and residents' health. Timely knowledge of the distribution of eggs from the two Pomacea spp. in a certain region can effectively reduce the cost of treatment and improve prevention effectiveness. Most of the existing methods are only able to identify eggs from the two Pomacea spp. or detect them in specific but not natural environments while they cannot achieve good results in the face of a complex real-world scene. This letter proposes a model for detecting eggs from the two Pomacea spp. based on dynamic convolution and multiscale feature fusion. The model can identify and locate the eggs of the two Pomacea spp. effectively. At the same time, we combined the proposed model with scale invariant feature transform (SIFT) algorithm to design a system for counting eggs of the two Pomacea spp., which can automatically identify the eggs in the actual natural environment and alleviate duplicate counting caused by image acquisition. Besides, we also built a dataset of 20,000 images of Pomacea canaliculata eggs and Pomacea maculata eggs from unmanned aerial vehicle (UAV) aerial photography. Experimental results showed that the proposed deep learning model has a better performance than others, and the proposed computer vision system can be successfully applied to support Pomacea spp. disease management.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000732167400001  
Indexed Date:  
2021-10-05  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Comp Sci & Technol, Nanjing 211106, Peoples R China  
Institution Address:  
MIIT Key Lab Pattern Anal & Machine Intelligence, Nanjing 211106, Peoples R China  
Institution:  
Institution:  
MIIT Key Lab Pattern Anal & Machine Intelligence

### 117. Multi-view Self-supervised Learning and Multi-scale Feature Fusion for

摘要: To address the challenges of the poor representation capability and low data utilization rate of end-to-end speech recognition models in deep learning, this study proposes an end-to-end speech recognition model based on multi-scale feature fusion and multi-view self-supervised learning (MM-ASR). It adopts a multi-task learning paradigm for training. The proposed method emphasizes the importance of inter-layer information within shared encoders, aiming to enhance the model's characterization capability via the multi-scale feature fusion module. Moreover, we apply multi-view self-supervised learning to effectively exploit data information. Our approach is rigorously evaluated on the Aishell-1 dataset and further validated its effectiveness on the English corpus WSJ. The experimental results demonstrate a noteworthy 4.6 % \documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\%$$\end{document} reduction in character error rate, indicating significantly improved speech recognition performance . These findings showcase the effectiveness and potential of our proposed MM-ASR model for end-to-end speech recognition tasks.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1370-4621  
eISSN:  
1573-773X  
入藏号:  
WOS:001216151500002  
Indexed Date:  
2024-05-13  
Institution Address:  
Beijing Univ Technol, Fac Informat Technol, Beijing, Peoples R China  
Institution:

### 118. An Automated Multi-scale Feature Fusion Network for Spine Fracture

摘要: Spine fractures represent a critical health concern with far-reaching implications for patient care and clinical decision-making. Accurate segmentation of spine fractures from medical images is a crucial task due to its location, shape, type, and severity. Addressing these challenges often requires the use of advanced machine learning and deep learning techniques. In this research, a novel multi-scale feature fusion deep learning model is proposed for the automated spine fracture segmentation using Computed Tomography (CT) to these challenges. The proposed model consists of six modules; Feature Fusion Module (FFM), Squeeze and Excitation (SEM), Atrous Spatial Pyramid Pooling (ASPP), Residual Convolution Block Attention Module (RCBAM), Residual Border Refinement Attention Block (RBRAB), and Local Position Residual Attention Block (LPRAB). These modules are used to apply multi-scale feature fusion, spatial feature extraction, channel-wise feature improvement, segmentation border results border refinement, and positional focus on the region of interest. After that, a decoder network is used to predict the fractured spine. The experimental results show that the proposed approach achieves better accuracy results in solving the above challenges and also performs well compared to the existing segmentation methods.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
被引频次合计:  
5  
ISSN:  
2948-2925  
eISSN:  
2948-2933  
入藏号:  
WOS:001204586200003  
PubMed ID:  
38622384  
Indexed Date:  
2024-04-24  
Institution Address:  
Cent South Univ, Sch Comp Sci & Engn, Changsha 410083, Hunan, Peoples R China  
Institution Address:  
Umm Al Qura Univ, Coll Comp & Informat Syst, Comp & Network Engn Dept, Mecca 24382, Saudi Arabia  
Institution:  
Institution:

### 119. An Insulator Defect Detection Model in Aerial Images Based on Multiscale

摘要: The faults caused by insulator defects will seriously threaten the operational safety of the power grid. Therefore, insulator defect detection plays a crucial role in inspecting transmission lines. Compared with traditional methods, the network such as You Only Look Once (YOLO) family based on deep learning has high accuracy and strong robustness in insulator recognition and fault detection. However, the performance of these network is usually affected by the shooting conditions as well as aerial images with diverse types of insulators and complex backgrounds, resulting in poor detection result. In addition, the relatively small insulator fault (bunch-drop) area in aerial images will make detection difficult. To solve these problems, this article proposes an improved insulator defect detection model based on YOLOv4 (ID-YOLO). To create our model, we design a new backbone network structure, cross-stage partial and residual split attention network (CSP-ResNeSt), that can solve the interference problem of complex backgrounds in aerial images to enhance the network's feature extraction capability. In addition, we adopt a new multiscale bidirectional feature pyramid network with simple attention module (Bi-SimAM-FPN), which can address the difficulty of identifying a small scale of insulator defects in an image for more efficient feature fusion. We experimentally demonstrate that the mean average precision (mAP) of the proposed model is 95.63%, which is 3.5% higher than that of the YOLOv4. Most importantly, the detection speed of this model can reach 63 frames/s, which meets the requirements of real-time detection of insulator bunch-drop faults.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
67  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
67  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000854588300004  
Indexed Date:  
2022-09-23  
Institution Address:  
Tianjin Chengjian Univ, Sch Comp & Informat Engn, Tianjin 300192, Peoples R China  
Institution Address:  
Tianjin Keyvia Elect Co Ltd, Tianjin 300384, Peoples R China  
Institution:  
Institution:  
Tianjin Keyvia Elect Co Ltd

### 120. Lightweight Target Detection Algorithm Based on Adaptive Spatial Feature

摘要: Aiming at the problems of complex network structure, difficult training and difficult deployment in mobile, and embedded devices of single-stage target detection in deep learning, a lightweight target detection algorithm based on adaptive spatial feature fusion is proposed. The proposed algorithm takes YOLOv4 as the basic framework of the network and uses lightweight MobileNet as the feature extraction network to reduce the network depth and training difficulty and improve the detection speed; an adaptive spatial feature fusion (ASFF) method is used to improve the poor effect of PANet on multi- scale feature fusion; by adding the output dimension of the network, the Gaussian algorithm is used to model the new dimension and output the uncertainty of the position of the prediction box; finally, the position loss function is redefined to improve the accuracy of position regression. The proposed algorithm takes the mask wearing detection robot during the epidemic as the deployment carrier to test the face mask wearing. The experimental results show that the detection accuracy of the proposed algorithm reaches 95. 92% and the detection speed reaches 19 frame/s. Compared with the original algorithm and other mainstream detection algorithms, the proposed algorithm is more suitable for deployment in mobile and embedded devices to realize real- time detection.  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1006-4125  
入藏号:  
WOS:000823174500031  
Indexed Date:  
2022-02-01  
Institution Address:  
Hunan Univ Sci & Technol, Sch Informat & Elect Engn, Xiangtan 411100, Hunan, Peoples R China  
Institution:

### 121. MCANet: A Multidimensional Channel Attention Residual Neural Network for

摘要: In the remote sensing image fusion field, fusion methods based on deep learning (DL) are the latest techniques in panchromatic sharpening (pansharpening). However, existing pansharpening methods based on neural networks cannot adequately inject the spatial feature information of panchromatic (PAN) images into fusion images, and they do not exploit the feature relationships between spatial locations, such as rows and columns of feature maps. To solve these problems, a multidimensional channel attention residual neural network (MCANet) is proposed in this article. To preserve the structural information in PAN images, a two-stream detail injection (TSDI) module is proposed, and the local skip connection operation is adopted to mine more spectral and structural information. A multidimensional channel attention (MCA) module is also designed to enable the network to learn the nonlinear mapping relationships between image spatial locations. In addition, a multiscale feature fusion (MSFF) module is designed to improve feature representation in the image fusion process, which is conducive to improving the pansharpening effect. The experimental results on the WorldView-2 (WV-2), GaoFen-2 (GF-2), and QuickBird (QB) datasets demonstrate that the proposed method outperforms state-of-the-art methods both visually and quantitatively.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000861485000014  
Indexed Date:  
2022-10-13  
Institution Address:  
Chongqing Univ Posts & Telecommun, Chongqing Key Lab Image Cognit, Chongqing 400065, Peoples R China  
Institution:

### 122. A Multiscale and Multidirection Feature Fusion Network for Road

摘要: The completeness of road extraction is very important for road application. However, existing deep learning (DP) methods of extraction often generate fragmented results. The prime reason is that DP-based road extraction methods use square kernel convolution, which is challenging to learn long-range contextual relationships of roads. The road often produce fractures in the local interference area. Besides, the quality of extraction results will be subjected to the resolution of remote sensing (RS) image. Generally, an algorithm will produce worse fragmentation when the used data differ from the resolution of the training set. To address these issues, we propose a novel road extraction framework for RS images, named the multiscale and multidirection feature fusion network (MSMDFF-Net). This framework comprises three main components: the multidirectional feature fusion (MDFF) initial block, the multiscale residual (MSR) encoder, and the multidirectional combined fusion (MDCF) decoder. First, according to the road's morphological characteristics, we develop a strip convolution module with a direction parameter (SCM-D). Then, to make the extracted result more complete, four SCM-D with different directions are used to MDFF-initial block and multidirectional combined fusion decoder (MDCF-decoder). Finally, we incorporate an additional branch into the residual network (ResNet) encoding module to build multiscale residual encoder (MSR-encoder) for improving the generalization of the model on different resolution RS image. Extensive experiments on three popular datasets with different resolution (Massachusetts, DeepGlobe, and SpaceNet datasets) show that the proposed MSMDFF-Net achieves new state-of-the-art results. The code will be available at https://github.com/wycloveinfall/MSMDFF-NET.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001196725900022  
Indexed Date:  
2024-04-18  
Institution Address:  
Univ Elect Sci & Technol China, Sch Automat Engn, Qingshuihe Campus, Chengdu 611731, Sichuan, Peoples R China  
Institution:

### 123. Efficient Defect Classification Using Few-Shot Image Generation and

摘要: Although deep learning has been proven to significantly outperform most traditional methods in the classification of large-scale balanced image datasets, collecting enough samples for defect classification is extremely time-consuming and costly. In this paper, we propose a lightweight defect classification method based on few-shot image generation and self-attention fused convolution features. We constructed a 4-class dataset using welding seam images collected from a solar cell module packaging production line. To address the issue of limited defect samples, especially for classes with less than 10 images, we implemented two strategies. Geometric enhancement techniques were first used to extend the defective images. Secondly, multi-scale feature fusion Generative Adversarial Networks (GANs) were utilized to further enhance the dataset. We then performed the feature-level fusion of convolution neural networks and self-attention networks, achieving a classification accuracy of 98.19%. Our experimental results demonstrate that the proposed model performs well in small sample defect classification tasks. And, it can be effectively applied to product quality inspection tasks in industrial production lines.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001254717700001  
Indexed Date:  
2024-07-06  
Institution Address:  
Univ Sci & Technol Beijing, Natl Ctr Mat Serv Safety, Beijing 100083, Peoples R China  
Institution Address:  
Southern Marine Sci & Engn Guangdong Lab, Zhuhai 519080, Peoples R China  
Institution:  
Institution:  
Southern Marine Sci & Engn Guangdong Lab

### 124. YOLO-OB: An improved anchor-free real-time multiscale colon polyp

摘要: The various sizes of polyps in colonoscopy images lead to low recall rates for current deep learning algorithms during polyp detection. To address this issue, we develop a new model called YOLO-OB, which significantly improves the recall rate for polyp detection while maintaining a very high precision. Our efforts mainly focus on the multiscale feature fusion structure and the polyp bounding box regression strategy for the detection head of YOLO-OB. Specifically, (1) we develop a bidirectional feature fusion structure, BiSPFPN, which efficiently facilitates feature fusion across different layers with varying receptive fields; (2) in contrast to existing polyp detection algorithms, which use specific constraints to filter candidate samples based on particular scale feature maps, we propose a novel polyp bounding box regression strategy, which matches candidate samples across all scales of feature maps by using the center point of the polyp ground truth bounding box. This strategy leverages multi-scale feature maps more effectively, enabling detecting polyps of different sizes on feature maps at any scale. Experiments on the public dataset SUN and the self-collected colon polyp dataset Union demonstrate that the proposed model significantly improves various performance metrics of polyp detection, especially the recall rate. Compared to the state-of-the-art results on the public dataset SUN, the proposed method achieves a 4.77% increase on recall rate from 93.46% to 98.23%. Furthermore, our model is able to achieve real-time polyp detection at a speed of 39 frames per second using a RTX3090 graphics card. The Python implementation of this work can be found here: https://github.com/seanyan62/YOLO-OB.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:001397671400001  
Indexed Date:  
2025-01-22  
Institution Address:  
Huazhong Univ Sci & Technol, Sch Comp Sci & Technol, Wuhan 430074, Hubei, Peoples R China  
Institution Address:  
Wuhan United Imaging Healthcare Surg Technol Co Lt, Intelligent Image Anal Dept, Wuhan, Hubei, Peoples R China  
Institution:  
Institution:  
Wuhan United Imaging Healthcare Surg Technol Co Lt

### 125. Defect Detection for Wear Debris Based on Few-Shot Contrastive Learning

摘要: In industrial defect detection tasks, the low probability of occurrence of severe industrial defects under normal production conditions has brought a great challenge for data-driven deep learning models that have just a few samples. Contrastive learning based on a sample pair makes it possible to obtain a great number of training samples and learn effective features quickly. In the field of industrial defect detection, the features of some defect instances have small category variance, and the scale of some defect instances has a great diversity. We propose a few-shot object detection network based on contrastive learning and multi-scale feature fusion. Aligned contrastive loss is adopted to increase the instance-level intra-class compactness and inter-class variance, and the misalignment problem is alleviated to a certain extent. A multi-scale fusion module is designed to recognize multi-scale defects by adaptively fusing features from different resolutions with the idea of exploiting the support branch's information. The robustness and efficiency of the proposed method were evaluated on an industrial wear debris defect dataset and the MS COCO dataset.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
2076-3417  
入藏号:  
WOS:000910811200001  
Indexed Date:  
2023-01-27  
Institution Address:  
Univ Sci & Technol Beijing, Sch Comp & Commun Engn, Beijing 100083, Peoples R China  
Institution:

### 126. Scattering Enhancement and Feature Fusion Network for Aircraft Detection

摘要: Aircraft detection in synthetic aperture radar (SAR) images is one challenging task due to the discreteness of aircraft scattering, the diversity of aircraft size, and the interference of background. In order to deal with these problems, a novel method named scattering enhancement and feature fusion network (SEFFNet) is here proposed to detect aircraft via combining traditional image processing and deep learning together. At first, a scattering information extraction and enhancement module (SIEEM) is proposed to highlight the scattering points of aircraft targets. Then, to more effectively focus on the location of aircraft targets, a space-to-depth coordinate attention module (SDCAM) is further designed, following which an efficient multi-scale feature fusion pyramid (FFP) is also introduced to fuse the semantic information of different layers. At last, a contextual fusion head (CFH) is built to improve the receptive field for better detecting aircraft. The experiments carried out on the popular datasets SADD and SAR-AIRcraft-1.0 show that SEFFNet is more appropriate for aircraft detection, especially the small-size aircraft detection, in comparison with other state-of-the-art (SOTA) methods. Taking the dataset SADD for example, on average, the precision, recall, F1-score, and APs values are respectively 2.8%, 2.6%, 2.7%, and 2.0% higher than the baseline network YOLOv5.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:001422045800034  
Indexed Date:  
2025-02-23  
Institution Address:  
Shanghai Jiao Tong Univ, Sch Sensing Sci & Engn, Shanghai Key Lab Intelligent Sensing & Recognit, Shanghai 200240, Peoples R China  
Institution Address:  
Natl Univ Def Technol, Coll Elect Sci & Technol, Changsha 410073, Peoples R China  
Institution Address:  
Tongji Univ, Ctr Digital Innovat, Shanghai 200092, Peoples R China  
Institution:  
Institution:  
Institution:

### 127. Image Reconstruction of Electrical Capacitance Tomography Based on

摘要: In order to solve the nonlinear ill-posed inverse problem in electrical capacitance tomography ( ECT),a multiscale dense connection network ( multi-scale densely connected network,MD-Net) model is proposed. The model consists of a multiscale feature fusion module and a densely connected block to further improve the reconstruction accuracy of images by fusing multiscale features. A flow-type data set is constructed by the MATLAB simulation experiment platform, and the learning and training of the training set are completed by using the nonlinear mapping ability of the densely connected network. The training effect is evaluated by using the test set. Static experiments are conducted on this basis. The simulation and static experiments results show that the method has the highest reconstruction accuracy,good noise immunity,and generalization ability compared with LBP,Landweber iterative algorithm,and other deep learning methods.  
摘要:  
为求解电容层析成像中的非线性病态反问题,提出了一种多尺度密集连接网络( MD-Net)模型。该模型由多尺度特征融合模块和密集连接块组成,以通过融合多尺度特征进一步提高图像的重建精度。通过MATLAB仿真实验平台构建了流型数据集,利用密集连接网络的非线性映射能力,完成训练集的学习与训练,并利用测试集进行训练效果评价。在此基础上进行了静态实验。仿真与静态实验结果均表明:与LBP、Landweber迭代算法和其他深度学习方法相比,该方法的重建精度最高、抗噪能力强,并具有良好的泛化能力。  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1000-1158  
入藏号:  
CSCD:7722247  
Indexed Date:  
2024-07-26  
Institution Address:  
(保定)华北电力大学自动化系, 保定, 河北 071003, 中国  
Institution Address:  
Department of Automation,North China Electric Power University(Baoding), Baoding, Hebei 071003, China  
Institution:  
(保定)华北电力大学自动化系  
Institution:  
Department of Automation,North China Electric Power University(Baoding)

### 128. Remaining useful life prediction towards cycling stability of organic

摘要: Organic electrochemical transistors (OECTs) show abundant potential in biosensors, artificial neuromorphic systems, brain-machine interfaces, etc With the fast development of novel functional materials and new device structures, OECTs with high transconductance (g(m) > mS) and good cycling stabilities (> 10,000 cycles) have been developed. While stability characterization is always time-consuming, to accelerate the development and commercialization of OECTs, tools for stability prediction are urgently needed. In this paper, OECTs with good cycling stabilities are realized by minimizing the gate voltage amplitude during cycling, while a remaining useful life (RUL) prediction framework for OECTs is proposed. Specifically, OECTs based on p(g2T-T) show tremendously enhanced stability which exhibits only 46.1% on-current (I-ON) and 33.2% peak g(m) decreases after 80,000 cycles (53 min). Then, RUL prediction is proposed based on the run-to-failure (RtF) aging tests (cycling stability test of OECTs). By selecting two aging parameters (I-ON and peak g(m)) as health indicators (HI), a novel multi-scale feature fusion (MFF) method for RUL prediction is proposed, which consists of a long short-term memory (LSTM) neural network based multi-scale feature generator (MFG) module for feature extraction and an attention-based feature fusion (AFF) module for feature fusion. Consequently, richer effective information is utilized to improve the prediction performance, where the experimental results show the superiority of the proposed framework on multiple OECTs in RUL prediction tasks. Therefore, by introducing such a powerful framework for the evaluation of the lifetime of OECTs, further optimization of materials, devices, and integrated systems relevant to OECTs will be stimulated. Moreover, this tool can also be extended to other relevant bioelectronics.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2053-1591  
入藏号:  
WOS:001154861600001  
Indexed Date:  
2024-02-09  
Institution Address:  
Univ Elect Sci & Technol China, Sch Automat Engn, Chengdu, Peoples R China  
Institution Address:  
Univ Elect Sci & Technol China, Sch Comp Sci & Engn, Chengdu, Peoples R China  
Institution:  
Institution:

### 129. SAC-RSM: A High-Performance UAV-Side Road Surveillance Model Based on

摘要: Efficiently and precisely identifying small items on traffic highways using unmanned aerial vehicle (UAV) platforms with limited resources is a crucial yet challenging job. This research suggests a speedy, accurate, and component-optimized road surveillance model (SAC-RSM) for UAVs. This model addresses slow detection speed, limited detection of small objects, and deployment difficulties. First, we designed a super-resolution-assisted learning branch in the network to balance the model's detection speed and accuracy. This branch learns the feature representation from low to high resolution. This branch uses multiscale feature fusion in the encoding stage to enhance the feature representation of small objects, thereby enhancing their detection accuracy. Second, to avoid the problem of cross-layer convolution, which results in the loss of fine-grained information and low-learning efficiency, we propose using the convolution-to-space-convolution (CSPC) module in the backbone network to improve model detection's robustness. Third, to achieve real-time detection, we realized the model using the Huawei Ascend compute architecture for neural networks (CANNs) framework to enable automatic quantization and parallel inference acceleration. Finally, we deployed the accelerated model to the embedded platform Atlas 200I developer kit (DK) A2. Compared to the baseline model, the proposed method shows significant increases in mean average precision (mAP) values for the VisDrone and DroneVehicle data sets, with increases of 17.4% and 9.4%, respectively. The proposed method achieves frames/s (FPS) of 38.3, which is 2.1 times faster than the baseline model, meeting the requirement for high-performance real-time detection in a UAV environment.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2327-4662  
入藏号:  
WOS:001351916500099  
Indexed Date:  
2024-11-19  
Institution Address:  
Nanjing Univ Posts & Telecommun, Sch Comp Sci, Nanjing 210023, Jiangsu, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Jiangsu High Technol Res Key Lab Wireless Sensor N, Nanjing 210023, Jiangsu, Peoples R China  
Institution Address:  
Yancheng Inst Technol, Sch Informat Engn, Yancheng 224000, Jiangsu, Peoples R China  
Institution:  
Institution:  
Institution:

### 130. MLFEU-NET: A Multi-scale Low-level Feature Enhancement Unet for breast

摘要: Breast lesions constantly threaten the health of females. Segmentation methods of breast lesions are important for clinical diagnosis, and neural networks have been widely used and played a significant role in this field. However, false detections and miss detections still exist due to the variability in the shape and location of breast lesions, artifacts and noise in breast ultrasound (BUS) images, and structural defects in conventional segmentation networks. In this paper, a multi-scale low-level feature enhancement Unet (MLFEU-net) structure is presented, consisting of the U-net structure, low-level feature enhancement block (LFEB), and a parallel multiscale feature fusion (PMFF) module. Specifically, LFEB enhances the detail information during shallow downsampling and further feature selection. Meanwhile, in the neck of Unet, PMFF module uses different scales of dilation convolution to provide different sizes of sensory fields, and to efficiently merge shallow and deep features filtered, leads to more accurate segmentation results. To evaluate the MLFEU-net's segmentation ability, five quantitative metrics were used on two breast ultrasound datasets, and it was compared with several advanced segmentation techniques. The results illustrate that our approach surpasses other methods in performance. Moreover, robustness experiments validate the effectiveness of our approach in achieving robust segmentation of breast lesions.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:001327841400001  
Indexed Date:  
2024-10-12  
Institution Address:  
Cent South Univ Forestry & Technol, Coll Elect Informat & Phys, Changsha 410004, Peoples R China  
Institution:

### 131. Face Detection Based on DF-Net

摘要: Face data have found increasingly widespread applications in daily life. To efficiently and accurately extract face information from input images, this paper presents a DF-Net-based face detection approach. A lightweight facial feature extraction neural network based on the MobileNet-v2 architecture is designed and implemented. By incorporating multi-scale feature fusion and spatial pyramid modules, the system achieves face localization and extraction across multiple scales. The proposed network is trained on the open-source face detection dataset WiderFace. The hyperparameters such as bottleneck coefficients and quality factors are discussed. Comparative experiments with other commonly used networks are carried out in terms of network model size, processing speed, and network extraction accuracy. Experimental results affirm the efficacy and robustness of this method, especially in challenging facial poses.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001083040900001  
Indexed Date:  
2023-10-28  
Institution Address:  
Shenzhen Univ, Coll Phys & Optoelect Engn, Key Lab Optoelect Devices & Syst, Minist Educ & Guangdong Prov, Shenzhen 518060, Peoples R China  
Institution:

### 132. MMS-EF: A Multi-Scale Modular Extraction Framework for Enhancing Deep

摘要: The analysis of land cover using deep learning techniques plays a pivotal role in understanding land use dynamics, which is crucial for land management, urban planning, and cartography. However, due to the complexity of remote sensing images, deep learning models face practical challenges in the preprocessing stage, such as incomplete extraction of large-scale geographic features, loss of fine details, and misalignment issues in image stitching. To address these issues, this paper introduces the Multi-Scale Modular Extraction Framework (MMS-EF) specifically designed to enhance deep learning models in remote sensing applications. The framework incorporates three key components: (1) a multiscale overlapping segmentation module that captures comprehensive geographical information through multi-channel and multiscale processing, ensuring the integrity of large-scale features; (2) a multiscale feature fusion module that integrates local and global features, facilitating seamless image stitching and improving classification accuracy; and (3) a detail enhancement module that refines the extraction of small-scale features, enriching the semantic information of the imagery. Extensive experiments were conducted across various deep learning models, and the framework was validated on two public datasets. The results demonstrate that the proposed approach effectively mitigates the limitations of traditional preprocessing methods, significantly improving feature extraction accuracy and exhibiting strong adaptability across different datasets.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2073-445X  
入藏号:  
WOS:001365596000001  
Indexed Date:  
2024-12-05  
Institution Address:  
Liaoning Tech Univ, Sch Mapping & Geosci, Fuxin 123000, Peoples R China  
Institution Address:  
Liaoning Tech Univ, Collaborat Innovat Inst Geospatial Informat Serv, Fuxin 123000, Peoples R China  
Institution:  
Institution:

### 133. A Segmentation Algorithm of Colonoscopy Images Based on Multi-Scale

摘要: Colorectal cancer is a common malignant tumor. Colorectal cancer is primarily caused by the cancerization of an adenomatous polyp. Segmentation of polyps in computer-assisted enteroscopy images is helpful for doctors to diagnose and treat the disease accurately. In this study, a segmentation algorithm of colonoscopy images based on multi-scale feature fusion is proposed. The proposed algorithm adopts ResNet50 as the backbone network to extract features. The shallow features are processed using the cross extraction module, thus increasing the receptive field, retaining the texture information, and fusing the processed shallow features and deep features at different proportions based on a multi-proportion fusion module. The proposed algorithm is capable of suppressing redundant information, removing background noise, and sharpening boundaries while acquiring considerable semantic information. As revealed by the results of the experiments on the published Kvasir-SEG dataset of intestinal polyps, the mean Dice coefficient and mean intersection over union were obtained as 0.9192 and 0.8873, better than that of existing mainstream algorithms. The result verifies the effectiveness of the proposed network and provides a reference for deep learning concerning the image processing and analysis of intestinal polyps.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
2079-9292  
入藏号:  
WOS:000845918300001  
Indexed Date:  
2022-09-04  
Institution Address:  
Anhui Univ, Sch Integrated Circuits, Hefei 230601, Peoples R China  
Institution Address:  
AnHui Engn Lab Agro Ecol Big Data, Hefei 230601, Peoples R China  
Institution:  
Institution:  
AnHui Engn Lab Agro Ecol Big Data

### 134. Multi-Scale Feature Fusion Attention Network for Building Extraction in

摘要: The efficient semantic segmentation of buildings in high spatial resolution remote sensing images is a technical prerequisite for land resource management, high-precision mapping, construction planning and other applications. Current building extraction methods based on deep learning can obtain high-level abstract features of images. However, the extraction of some occluded buildings is inaccurate, and as the network deepens, small-volume buildings are lost and edges are blurred. Therefore, we introduce a multi-resolution attention combination network, which employs a multiscale channel and spatial attention module (MCAM) to adaptively capture key features and eliminate irrelevant information, which improves the accuracy of building extraction. In addition, we present a layered residual connectivity module (LRCM) to enhance the expression of information at different scales through multi-level feature fusion, significantly improving the understanding of context and the capturing of fine edge details. Extensive experiments were conducted on the WHU aerial image dataset and the Massachusetts building dataset. Compared with state-of-the-art semantic segmentation methods, this network achieves better building extraction results in remote sensing images, proving the effectiveness of the method.  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
4  
ISSN:  
2079-9292  
入藏号:  
WOS:001182832300001  
Indexed Date:  
2024-03-31  
Institution Address:  
Zhengzhou Univ Light Ind, Sch Comp & Commun Engn, Zhengzhou 450002, Peoples R China  
Institution:

### 135. Densely integrated multi-branch attentive net for image dehazing

摘要: Single image dehazing is a fundamental but challenging task in image processing. Various deep learning-based methods have achieved great dehazing performance. However, there are still hazy residues, even color distortion and texture loss when removing haze from complex outdoor images in dense hazy scenes. A densely integrated multi-branch attentive net for image dehazing is proposed in the paper to address the above problems. The network includes a multi-scale feature extraction module and a dense-feature fusion module. The multi-scale feature extraction module adopts a multi-branch structure composed of residual channel attention blocks, which can expand the receptive field and filter the extracted features of diverse scales by weighting for fusion. It raises network learning accuracy. The dense-feature fusion module contains a multi-level feature fusion module for front and back layers, a color information renovation module, and a feature enhancement module. It achieves dynamically adjusting the channel weights of features at diverse scales, learning rich context information and suppressing redundant information, and compensating for the lack of color and texture information. The dense- feature fusion module bolsters the generalization capability of the network. The proposed method achieves superior objective and subjective evaluation results via quantitative and qualitative experiments on synthetic hazy images and natural hazy images, with better generalization ability and dehazing effect than the current SOTA dehazing algorithms, and effectively ameliorates the color distortion and incomplete dehazing.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2090-4479  
eISSN:  
2090-4495  
入藏号:  
WOS:001274236400001  
Indexed Date:  
2024-07-28  
Institution Address:  
Shandong Technol & Business Univ, Sch Comp Sci & Technol, Yantai 264005, Peoples R China  
Institution Address:  
Shandong Technol & Business Univ, Sch Informat & Elect Engn, Yantai 264005, Peoples R China  
Institution:  
Institution:

### 136. Anomaly Detection in Chest X-rays Based on Dual-Attention Mechanism and

摘要: The efficient and automatic detection of chest abnormalities is vital for the auxiliary diagnosis of medical images. Many studies utilize computer vision and deep learning approaches involving symmetry and asymmetry concepts to detect chest abnormalities, and achieve promising findings. However, an accurate instance-level and multi-label detection of abnormalities in chest X-rays remains a significant challenge. Here, a novel anomaly detection method for symmetric chest X-rays using dual-attention and multi-scale feature fusion is proposed. Three aspects of our method should be noted in comparison with the previous approaches. We improved the deep neural network with channel-dimensional and spatial-dimensional attention to capture the abundant contextual features. We then used an optimized multi-scale learning framework for feature fusion to adapt to the scale variation in the abnormalities. Considering the influence of the data imbalance and other factors, we introduced a seesaw loss function to flexibly adjust the sample weights and enhance the model learning efficiency. The rigorous experimental evaluation of a public chest X-ray dataset with fourteen different types of abnormalities demonstrates that our model has a mean average precision of 0.362 and outperforms existing methods.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2073-8994  
入藏号:  
WOS:000960029400001  
Indexed Date:  
2023-04-16  
Institution Address:  
Xiangnan Univ, Sch Comp & Artificial Intelligence, Chenzhou 423300, Peoples R China  
Institution Address:  
Xiangnan Univ, Hunan Engn Res Ctr Adv Embedded Comp & Intelligent, Chenzhou 423300, Peoples R China  
Institution Address:  
Xiangnan Univ, Key Lab Med Imaging & Artificial Intelligence Huna, Chenzhou 423300, Peoples R China  
Institution Address:  
Jilin Univ, Sch Comp Sci & Technol, Changchun 130012, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 137. DAN-SuperPoint: Self-Supervised Feature Point Detection Algorithm with

摘要: In view of the poor performance of traditional feature point detection methods in low-texture situations, we design a new self-supervised feature extraction network that can be applied to the visual odometer (VO) front-end feature extraction module based on the deep learning method. First, the network uses the feature pyramid structure to perform multi-scale feature fusion to obtain a feature map containing multi-scale information. Then, the feature map is passed through the position attention module and the channel attention module to obtain the feature dependency relationship of the spatial dimension and the channel dimension, respectively, and the weighted spatial feature map and the channel feature map are added element by element to enhance the feature representation. Finally, the weighted feature maps are trained for detectors and descriptors respectively. In addition, in order to improve the prediction accuracy of feature point locations and speed up the network convergence, we add a confidence loss term and a tolerance loss term to the loss functions of the detector and descriptor, respectively. The experiments show that our network achieves satisfactory performance under the Hpatches dataset and KITTI dataset, indicating the reliability of the network.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
eISSN:  
1424-8220  
入藏号:  
WOS:000768801700001  
PubMed ID:  
35271087  
Indexed Date:  
2022-03-23  
Institution Address:  
Shandong Univ Technol, Sch Mech Engn, Zibo 255000, Peoples R China  
Institution Address:  
Beijing Inst Technol, Sch Opt & Photon, Beijing 100081, Peoples R China  
Institution:  
Institution:

### 138. MicrosMobiNet: A Deep Lightweight Network With Hierarchical Feature

摘要: In recent advancements of lightweight deep architectures for edge devices, most of the works follow a typical MobileNet pipeline designed for computer vision tasks which is not very appropriate for microscopy image analysis. Certainly, the design of the dedicated lightweight network for highly complex microscopy image analysis has not been attempted so far. Therefore, this work proposes a new deep lightweight network, "MicrosMobiNet" having multiscale feature extraction mechanism for bright-field microscopy image analysis on a mobile-edge computing framework. It consists of three key attributes-depth-wise separable convolution for making the network lightweight, multiple kernels with hierarchical feature fusion to extract complex features, and residual connection to keep network deep. Experimental validations have been conducted by two different microscopy image data sets-plant (potato tuber) and histopathology (cancer cell) generated by two different image generation modalities. In the experiment, multiclass and multilabel classification tasks have been evaluated by measuring accuracy, F1-score, and error. In the ablation study, the key attributes of the network have been verified. The results and analysis show that the MicrosMobiNet can achieve classification accuracy up to 98.43% and 96.25% for plant and cancer cells with minimum error 8.38% and 10.03%, respectively. In a comparative study, the MicrosMobiNet outperforms the existing lightweight state-of-the-art methods with fewer parameters (1.9M) and FLOPs count (42M). Finally, the new network has been implemented on an edge device, Smartphone (Android platform) which is working satisfactorily with high speed (140 ms) and very low memory (7.4 MB). Hence, the network exhibits its superiority in bright-field microscopy image analysis on mobile-edge computing platforms in a lightweight deep learning framework.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
ISSN:  
2327-4662  
入藏号:  
WOS:001203463700064  
Indexed Date:  
2024-06-25  
Institution Address:  
Indian Inst Informat Technol Guwahati, Dept Elect & Commun Engn, Gauhati 781015, India  
Institution:  
Indian Inst Informat Technol Guwahati

### 139. FFNet: Feature Fusion Network for Few-shot Semantic Segmentation

摘要: Semantic segmentation aims at assigning a category label to each pixel in an image. Deep neural networks have achieved many breakthrough research achievements on this task. Nevertheless, there exist two critical bottleneck problems to be solved. First, deep neural networks usually need to be trained on large-scale labeled datasets, which are expensive to obtain or label. Second, traditional semantic segmentation methods are difficult to predict unseen classes after training. To address these problems, few-shot semantic segmentation is proposed, and recent methods have achieved impressive performance. However, many of the existing approaches ignore the semantic correlation between data and fail to generate discriminative features for the semantic segmentation. In this paper, to address the above issue, we propose a feature fusion network (FFNet) for few-shot semantic segmentation to enhance the discriminative ability of the learned data representations. Specifically, a task attention module is devised to learn the semantic correlation between data. Then, a multi-scale feature fusion module is trained to adaptively fuse the contextual information at multiple scale, thus capturing multi-scale object information. To the end, the proposed FFNet experiments conducted on the PASCAL-5(i) and COCO-20(i) datasets demonstrate the superiority of our proposed FFNet and show its advantage over existing approaches.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
16  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
17  
ISSN:  
1866-9956  
eISSN:  
1866-9964  
入藏号:  
WOS:000745591500003  
Indexed Date:  
2022-01-29  
Institution Address:  
Chinese Acad Sci, Inst Informat Engn, Beijing 100000, Peoples R China  
Institution Address:  
Ocean Univ China, Inst Artificial Intelligence, Lab Intelligent Data Anal & Syst IDAS, Qingdao 266100, Peoples R China  
Institution Address:  
Ocean Univ China, Coll Comp Sci & Technol, Qingdao 266100, Peoples R China  
Institution:  
Institution:  
Institution:

### 140. Deep learning and IoT enabled digital twin framework for monitoring

摘要: Early detection of cracks enables timely mitigation and maintenance actions, ensuring the safety of personnel and equipment within the open-pit coal mine. Monitoring open-pit coal mines and cracks is essential for the safety of workers and for saving national assets. Digital twins (DTs) can be crucial in open-pit coal mine crack detection. DTs enable continuous real-time monitoring of the open-pit mine, including its structures and surrounding environment. Various sensors and internet-of-things devices can be deployed to collect data on factors such as ground movement and strain. Integrating this data into the DT makes it possible to identify and analyze anomalous behavior or changes that may indicate crack formation or propagation. Deep learning-based networks are a crucial factor in detecting open-pit coal mine cracks. In this work, we propose a deep learning-based densely connected lightweight network incorporated into the DT-based framework for detecting cracks and taking predictive maintenance-based decisions by combining historical data, real-time sensor data, and predictive models. The proposed DT-based framework provides insights into the potential crack formation, allowing for proactive maintenance and mitigation measures. We compare the performance of our proposed network on different evaluation measures such as precision, recall, overall accuracy, mean average precision, F1-score, and kappa coefficient, where our proposed lightweight multiscale feature fusion-based network outperformed all other state-of-the-art deep neural networks. We also achieved the best performance on mean average precision by surpassing all other models. Additionally, we also compared the performance of our proposed network with U-Net and recurrent neural network on model training and prediction time benchmarks by outperforming those cutting-edge models.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
2296-598X  
入藏号:  
WOS:001094992000001  
Indexed Date:  
2023-11-26  
Institution Address:  
China Coal Huajin Grp Co Ltd, Yuncheng, Shaanxi, Peoples R China  
Institution:  
China Coal Huajin Grp Co Ltd

### 141. An Automatic Deep Segmentation Network for Pixel-Level Welding Defect

摘要: Accurate welding defect location is of great significance to modern manufacturing, which could be used for accurate quality evaluation and precise repairing decision-making basis of different products. Nevertheless, accurate welding defect location is still a challenging task due to some complex factors, such as complex backgrounds, low contrast, weak texture, and class imbalance issue. Recently, deep learning has got great development due to its strong feature expression ability, which has been widely applied into defect detection, but it still exists certain shortcomings on segmentation tasks with the class unbalanced issue or microdefects. To address these issues, with the encoder & x2013;decoder network architecture, a novel welding defect location method is proposed with an attention-guided segmentation network. To reduce the contextual information loss of the deep encoder module after multiple convolution and pooling operations, a multiscale feature fusion block is proposed to embed into a U-shaped network (U-Net) to acquire more information. On the basis, combined with a bidirectional convolutional long short-term memory (BiConvLSTM) block, an improved attention block is integrated into the skip connections between the encoder path and the decoder path to capture the global, long-range contexts and emphasize target regions, contributing to locate welding defect areas and enhance the segmentation ability on microdefects. Meanwhile, to address the foreground & x2013;background class imbalance issue, a hybrid loss function combined with binary cross-entropy (BCE) and loss functions is proposed to effectively utilize their unique excellent characteristics for accurate defect segmentation. Experiment results on the public <italic>GDXray</italic> dataset show that the proposed segmentation method could obtain a competitive segmentation performance compared with other advanced segmentation models.  
Web of Science 核心合集中的 "被引频次":  
54  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
55  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000766300200007  
Indexed Date:  
2022-03-31  
Institution Address:  
Zhengzhou Univ, Sch Elect Engn, Zhengzhou 450001, Henan, Peoples R China  
Institution Address:  
Zhengzhou Univ, Robot Percept & Control Engn Lab, Zhengzhou 450001, Henan, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Management & Control Complex Syst, Beijing 100190, Peoples R China  
Institution:  
Institution:  
Institution:

### 142. Lithography hotspot detection through multi-scale feature fusion

摘要: Lithography hotspot (LHS) detection is crucial for achieving manufacturability design in integrated circuits (ICs) and ensuring the final yield of ICs chips. Recognizing the challenges posed by conventional deep learning-based methods for lithographic hotspot detection in meeting the demands of advanced IC manufacturing accuracy, this study introduces an LHS detection approach. This approach leverages multiscale feature fusion to identify defects in lithographic layout hotspots accurately. This method incorporates the convolutional block attention module into the backbone network to enhance the focus of the model on the layout area. Additionally, a feature pyramid is employed to merge deep and shallow features from the layout pattern, significantly enhancing the capability of hotspot detection network to extract both image and semantic features. Concurrently, by utilizing a dense block that directly interconnects various layers, the network gains the capacity to capture the correlation between low-level and high-level features, thereby enhancing the perceptual capabilities of the model. Experimental results demonstrate the superiority of the algorithm across accuracy, false alarm, F1 score, and overall detection simulation time compared to alternative lithographic hotspot detection algorithms. (c) 2024 Society of Photo-Optical Instrumentation Engineers (SPIE)  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1932-5150  
eISSN:  
2708-8340  
入藏号:  
WOS:001208695800014  
Indexed Date:  
2024-05-31  
Institution Address:  
Anhui Univ Sci & Technol, Sch Comp Sci & Engn, Huainan, Peoples R China  
Institution Address:  
Anhui Univ Sci & Technol, Sch Artificial Intelligence, Huainan, Peoples R China  
Institution Address:  
Hefei Univ Technol, Sch Microelect, Hefei, Peoples R China  
Institution:  
Institution:  
Institution:

### 143. Automatic detection of three cell types in a microscope image based on

摘要: With the continuous integration of deep learning and the technique of molecular biology, target detection models must accurately detect the position of each cell in the image and classify it correctly. We present a model for the multi-scale feature fusion of the existing human cell image dataset based on Gaussian mixedly clustering. First, a novel feature extraction network for extracting preliminary features at picture multi scales was presented, which was based on a residual neural network with Instance Normalization and a Mish activation function. Second, the presented model adopts the idea of feature fusion and introduced a new type of feature fusion network to integrate feature graphs on different scales. Furthermore, a Gaussian hybrid clustering algorithm was proposed to cluster the hyperparameters. Based on the experimental results, the average accuracy of the proposed model in the human cell image dataset exceeds 0.96, which improves by 11.9% compared with the existing target detection methods in the same field. Experiments show that the proposed model had been adapted to datasets with uneven sample distribution, providing new ideas for research on medical images.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1864-063X  
eISSN:  
1864-0648  
入藏号:  
WOS:000842893100001  
PubMed ID:  
36054791  
Indexed Date:  
2022-08-27  
Institution Address:  
Shenyang Univ Chem Technol, Coll Comp Sci & Technol, Shenyang 110016, Peoples R China  
Institution:

### 144. Multi-scale hierarchical feature fusion network for change detection

摘要: Change Detection (CD) in Very High-Resolution (VHR) images can bring more detailed and valuable geoinformation for geographic information system. However, existing CD methods are still limited by the poor recognition of multi-scale land cover objects with varied shapes in VHR CD tasks. Besides, obtaining accurate change information in complex scenes is still challenging since high-frequency component tends to be suppressed in the most deep learning-based CD methods. To alleviate these problems, we propose a novel encoder-decoder network, Multi-scale Hierarchical Feature Fusion Network (MHF2Net), which utilizes all-scale feature fusion and global high-frequency enhancement strategies. The key functional modules are High-frequency Enhancement Blocks (HEBs) and a Layer-wise Multi-scale Feature Fusion Module (LMF2M). HEB provides auxiliary high-frequency information to further finely annotate the changed land cover. LMF2M fully integrates hierarchical features to better fuse semantic and spatial information at each layer, thus better detecting multi-scale changed objects. And in LMF2M, a Self-weighted Attention Block (SAB) is constructed to throttle uninformative features without supervised parameters. Asa result, MHF2Net is able to map the changed objects of varied scales and shapes with quite high accuracy. Extensive experiments are conducted over three public VHR change detection data sets. And the corresponding results suggest that the proposed method achieves state-of-the-art change detection performance in comparison with several recently proposed methods.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:001391067300001  
Indexed Date:  
2025-01-11  
Institution Address:  
Xidian Univ, Sch Elect Engn, Xian, Peoples R China  
Institution Address:  
Inner Mongolia Normal Univ, Coll Math Sci, Hohhot 010022, Peoples R China  
Institution Address:  
Swinburne Univ Technol, Dept Comp Technol, Hawthorn, Vic 3122, Australia  
Institution Address:  
Shaanxi Univ Sci & Technol, Sch Elect Informat & Artificial Intelligence, Xian, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 145. A Novel Lung Nodule Detection and Recognition Model Based on Deep

摘要: To solve the problems of missing and false detection of pulmonary nodules in complex lung environments, as well as trivial and inefficient detection procedures, an end-to-end pulmonary nodules detection and recognition model based on deep learning was proposed. Innovation and improvement are made on the basis of YOLOv5. In the feature extraction stage of the model, a convolutional structure integrating self-attention mechanism is proposed to capture the global feature and the dependence relationship of long-distance information, and screen the key pathological information. Then, a convolution structure integrating internal convolution operators is proposed to reduce the computational redundancy in the feature channel and improve the inference speed of the model. In the feature fusion stage of the model, the structure of cross-scale coordinate attention feature fusion is proposed, and the different features enhanced with attention are weighted by jumping links to promote the fusion of multi-scale feature information. The proposed model obtained 97.8% mAP@0.5 indexes in the self-built diagnosis and treatment data set of pulmonary nodules in Huaihai area. The pulmonary nodule detection model proposed in this paper can significantly reduce the false positive rate and obtain the location and classification results of diseased nodules with higher detection accuracy and faster detection speed, which has important practical value in clinical application.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001346089600001  
Indexed Date:  
2024-11-11  
Institution Address:  
Xuzhou 1 Peoples Hosp, Xuzhou 221002, Jiangsu, Peoples R China  
Institution Address:  
China Univ Min & Technol, Sch Informat & Control Engn, Xuzhou 221116, Jiangsu, Peoples R China  
Institution:  
Xuzhou 1 Peoples Hosp  
Institution:

### 146. YOLO-Lite: An Efficient Lightweight Network for SAR Ship Detection

摘要: Automatic ship detection in SAR images plays an essential role in both military and civilian fields. However, most of the existing deep learning detection methods introduce complex models and huge calculations while improving the detection accuracy, which is not conducive to the application of real-time ship detection. To solve this problem, an efficient lightweight network YOLO-Lite is proposed for SAR ship detection in this paper. First, a lightweight feature enhancement backbone (LFEBNet) is designed to reduce the amount of calculation. Additionally, a channel and position enhancement attention (CPEA) module is constructed and embedded into the backbone network to more accurately locate the target location by capturing the positional information. Second, an enhanced spatial pyramid pooling (EnSPP) module is customized to enhance the expression ability of features and address the position information loss of small SAR ships in high-level features. Third, we construct an effective multi-scale feature fusion network (MFFNet) with two feature fusion channels to obtain feature maps with more position and semantic information. Furthermore, a novel confidence loss function is proposed to effectively improve the SAR ship target detection accuracy. Extensive experiments on SSDD and SAR ship datasets verify the effectiveness of our YOLO-Lite, which can not only accurately detect SAR ships in different backgrounds but can also realize a lightweight architecture with low computation cost.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
38  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
39  
eISSN:  
2072-4292  
入藏号:  
WOS:001045608900001  
Indexed Date:  
2023-08-23  
Institution Address:  
Henan Univ Technol, Sch Artificial Intelligence & Big Data, Zhengzhou 450001, Peoples R China  
Institution Address:  
Henan Univ Technol, Coll Informat Sci & Engn, Key Lab Grain Informat Proc & Control, Minist Educ, Zhengzhou 450001, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100190, Peoples R China  
Institution:  
Institution:  
Institution:

### 147. Road Extraction Using a Dual Attention Dilated-LinkNet Based on

摘要: Automatic extraction of road from multisource remote sensing data has always been a challenging task. Factors such as shadow occlusion and multisource data alignment errors prevent current deep learning-based road extraction methods from acquiring road features with high complementarity, redundancy, and crossover. Unlike previous works that capture contexts by multiscale feature fusion, we propose a dual attention dilated-LinkNet (DAD-LinkNet) to adaptively integrate local road features with their global dependencies by joint using satellite image and floating vehicle trajectory data. First, a joint least-squares feature matching-based floating vehicle trajectory correction model is used to correct the floating vehicle trajectory; then a convolutional network model DAD-LinkNet based on a dual-attention mechanism is proposed, and road features are extracted from the channel domain and spatial domain of the target image in turn by constructing a dual-attention module in the dilated convolutional layer and adopting a cascade connection; a weighted hyperparameter loss function is used as the loss function of the model; finally, the road extraction is completed based on the proposed DAD-LinkNet model. Experiments on three datasets show that the proposed DAD-LinkNet model outperforms the state-of-the-art methods in terms of accuracy and connectivity.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
17  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
18  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000711641000016  
Indexed Date:  
2021-11-11  
Institution Address:  
Northwestern Polytech Univ, Sch Software, Xian 710072, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Shenzhen Res Inst, Shenzhen 518057, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Remote Sensing & Informat Engn, Wuhan 430079, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Dept Land Surveying & Geoinformat, Hong Kong, Peoples R China  
Institution Address:  
Xian Univ Technol, Sch Comp Sci & Engn, Xian 710048, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Sch Automat, Xian 710072, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 148. SPSNet: A Selected Pyramidal Shape-Constrained Network for SAR Target

摘要: The complex background and coherent speckle noise in synthetic aperture radar (SAR) images presents a significant challenge for the detection and recognition of SAR small targets. For deep neural networks, the robust feature learning method and effective loss function could enhance the accuracy of SAR target detection and reduce false alarm rates. However, many of feature enhancement networks based on feature pyramid network (FPN) have limited ability to capture feature interaction between different branches. In addition, the design of loss function cannot generate samples that better match the shape of SAR targets for network training. In this article, we propose a selected pyramidal shape-constrained network (SPSNet) to alleviate these problems. A feature fusion paradigm, including a spatial selection block and a dynamic channel attention module, are inserted into FPN for adaptive multiscale feature selection and feature enhancement in spatial-channel feature dimension. Both of these modules could capture the distinguishable features of SAR targets. Furthermore, the shape information of SAR target is utilized into detection loss to enhance the quality of SAR detection box sampling points in a soft threshold style, thereby enhancing the model's detection for SAR targets. The experimental results of three challenge SAR detection datasets illustrate that SPSNet gains superior performances.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001282396300005  
Indexed Date:  
2024-08-20  
Institution Address:  
Nanjing Univ Posts & Telecommun, Sch Comp Sci, Nanjing 210023, Peoples R China  
Institution Address:  
Jiangsu Prov Engn Res Ctr Airborne Detecting & In, Nanjing 210049, Peoples R China  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Key Lab Radar Imaging & Microwave Photon, Minist Educ, Nanjing 211106, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Coll Automat, Nanjing 210023, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Coll Artificial Intelligence, Nanjing 210023, Peoples R China  
Institution:  
Institution:  
Jiangsu Prov Engn Res Ctr Airborne Detecting & In  
Institution:  
Institution:  
Institution:

### 149. Maize leaf disease identification based on WG-MARNet

摘要: In deep learning-based maize leaf disease detection, a maize disease identification method called Network based on wavelet threshold-guided bilateral filtering, multi-channel ResNet, and attenuation factor (WG-MARNet) is proposed. This method can solve the problems of noise, background interference, and low detection accuracy of maize leaf disease images. To begin, a processing layer called Wavelet threshold guided bilateral filtering (WT-GBF) based on the WG-MARNet model is employed to reduce image noise and perform high and low-frequency decomposition of the input image using WT-GBF. This increases the input image's resistance to environmental interference and feature extraction capability. Secondly, for the multiscale feature fusion technique, an average down-sampling and tiling method is employed to increase feature representation and limit the risk of overfitting. Then, on high and low-frequency multi-channel, an attenuation factor is introduced to optimize the performance instability during training of the deep network. Finally, when the convergence and accuracy are compared, PRelu and Adabound are used instead of the Relu activation function and the Adam optimizer. The experimental results revealed that our method's average recognition accuracy was 97.96%, and the detection time for a single image was 0.278 seconds. The average detection accuracy has been increased. The method lays the groundwork for the precise control of maize diseases in the field.  
Web of Science 核心合集中的 "被引频次":  
8  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
5  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
8  
ISSN:  
1932-6203  
入藏号:  
WOS:000794035700096  
PubMed ID:  
35483023  
Indexed Date:  
2022-06-01  
Institution Address:  
Cent South Univ Forestry & Technol, Coll Comp & Informat Engn, Changsha, Hunan, Peoples R China  
Institution Address:  
Acad Agr Sci, Plant Protect Res Inst, Changsha, Hunan, Peoples R China  
Institution Address:  
Natl Univ Def Technol, Changsha, Hunan, Peoples R China  
Institution:  
Institution:  
Acad Agr Sci  
Institution:

### 150. Electrical load forecasting based on the fusion of multi-scale features

摘要: Currently, deep learning methods have become prevalent in the field of electrical load forecasting. These approaches have shown a great potential to map complex nonlinear feature interactions. However, many existing electrical load forecasting models based solely on deep learning suffer from various limitations including the inability to perceive and integrate multi-scale features, the absence of continuous information of electrical load series and capturing fine-grained and hidden temporal pattern of electrical load series. In order to address these issues, in this paper we propose an improved model based on neural ordinary differential equations (NODEs), which possesses the ability of adaptive fusion, multi-scale feature perception, and representation. This model strengthens the effective decomposition of multi-scale features with the NODE-series block and enhances the multi-scale feature extraction and fusion ability with the NODE-split block. The experimental results show that the proposed model outperforms five baseline models and three block ablation experiments prove the necessity of the blocks.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0920-8542  
eISSN:  
1573-0484  
入藏号:  
WOS:001337777100009  
Indexed Date:  
2024-10-26  
Institution Address:  
Lanzhou Univ, Sch Informat Sci & Technol, Lanzhou, Peoples R China  
Institution Address:  
Univ Wollongong, Sch Comp & Informat Technol, Wollongong, Australia  
Institution Address:  
Univ Technol Sydney, Sch Comp Sci, Sydney, Australia  
Institution:  
Institution:  
Institution:

### 151. Small-Object Detection Based on YOLO and Dense Block via Image

摘要: Small-object detection is a basic and challenging problem in computer vision tasks. It is widely used in pedestrian detection, traffic sign detection, and other fields. This paper proposes a deep learning small-object detection method based on image super-resolution to improve the speed and accuracy of small-object detection. First, we add a feature texture transfer (FTT) module at the input end to improve the image resolution at this end as well as to remove the noise in the image. Then, in the backbone network, using the Darknet53 framework, we use dense blocks to replace residual blocks to reduce the number of network structure parameters to avoid unnecessary calculations. Then, to make full use of the features of small targets in the image, the neck uses a combination of SPPnet and PANnet to complete this part of the multi-scale feature fusion work. Finally, the problem of image background and foreground imbalance is solved by adding the foreground and background balance loss function to the YOLOv4 loss function part. The results of the experiment conducted using our self-built dataset show that the proposed method has higher accuracy and speed compared with the currently available small-target detection methods.  
在中国科学引文数据库中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
32  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
34  
ISSN:  
2169-3536  
入藏号:  
WOS:000641951900001  
Indexed Date:  
2021-04-25  
Institution Address:  
Yangtze Univ, Sch Elect Informat, Jingzhou 434023, Peoples R China  
Institution Address:  
Yangtze Univ, Natl Demonstrat Ctr Expt Elect & Elect Educ, Jingzhou 434023, Peoples R China  
Institution Address:  
Yangtze Univ, Western Inst, Karamay 834000, Peoples R China  
Institution Address:  
Yangtze Univ, Sch Comp Sci, Jingzhou 434023, Peoples R China  
Institution Address:  
Cent South Univ, Sch Comp Sci & Engn, Changsha 410083, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 152. Accurate and lightweight cloud detection method based on cloud and snow

摘要: Cloud detection is a critical stage in remote sensing image preprocessing. However, when there is snow on the underlying surface of scenes,the general cloud detection methods wouldbe easily affected. As a result,the cloud detection accuracy of these methods would reduce.Furthermore,most available cloud detection datasets are of medium-resolution and do not focus on the cloud and snow coexistence study areas. As a result, a cloud detection dataset has been created and released based on high-resolution cloud-snow coexistence remote sensing images. Meanwhile, this study suggests a convolution neural network termed RDC-Net for cloud detection in high-resolution cloud and snow coexistence images. The RDC-Net contains the reconstructible multiscale feature fusion module for multiscale cloud feature extraction, the dual adaptive feature fusion module for effective cloud feature representation reconstruction,and the controllably deep gradient guidance flows module for unbiased network gradient descent guidance. Benefiting from the above technical components, the network can enhance the robustness of cloud detection in complicated regions and facilitate lightweight deployment of the network. The experimental results show that the RDC-Net has an excellent anti-interference capacity for highlighted ground objects and has outstanding detection performance for thin clouds and clouds over snow. Furthermore,the RDC-Net has fewer parameters and floating-point operations, making it acceptable for industrial production and application.  
摘要:  
云检测是遥感图像预处理过程中的关键步骤,但是当场景的下垫面中存在雪时,常用的云检测方法易受到干扰而降低云检测精度。此外,现有云检测数据集多为中等分辨率,且并未强调探讨云雪共存区域。因此,本文创建发布了基于高分辨率云雪共存遥感影像的云检测数据集CloudS,并提出了一种面向高分辨率云雪共存场景的轻量云检测卷积神经网络RDC-Net。RDC-Net中包含可重构多尺度特征融合模块以用于多尺度云特征的提取;双重自适应特征融合模块以对有效云特征实现表征重建;可控深层梯度指导流模块进行网络梯度下降的无偏指导。受益于上述几个技术组件,该网络能进一步提升复杂区域云检测的稳健性并促进部署的轻量化。试验结果表明,本文方法对遥感影像中的薄云及雪域上空的云具有极佳的提取能力,同时对雪等高亮地物具有良好的抗干扰能力。此外,RDC-Net具有极少的参数量与前向推理浮点运算量,这也使得其适合于实际的工业生产部署。  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1001-1595  
入藏号:  
CSCD:7403171  
Indexed Date:  
2023-06-02  
Institution Address:  
长江大学地球科学学院, 武汉, 湖北 430100, 中国  
Institution Address:  
长江大学地球科学学院;;自然资源部环鄱阳湖区域矿山环境监测与治理重点实验室, ;;自然资源部环鄱阳湖区域矿山环境监测与治理重点实验室, 武汉;;南昌, 湖北;;江西 430100;;330013, 中国  
Institution Address:  
长江大学地球科学学院;;湖南科技大学;;城市空间信息工程北京市重点实验室, ;;测绘遥感信息工程湖南省重点实验室;;城市空间信息工程北京市重点实验室, 武汉;;湘潭;;, 湖北;;湖南;;北京 430100;;411201;;100045, 中国  
Institution Address:  
School of Geosciences, Yangtze University, Wuhan, Hubei 430100, China  
Institution Address:  
School of Geosciences, Yangtze University;;Key Laboratory of Mine Environmental Monitoring and Improving around Poyang Lake, Ministry of Natural Resources, ;;Key Laboratory of Mine Environmental Monitoring and Improving around Poyang Lake, Ministry of Natural Resources, Wuhan;;Nanchang, ;; 430100;;330013  
Institution Address:  
School of Geosciences, Yangtze University;;Mapping and Remote Sensing, Hunan University of Science and Technology;;Beijing Key Laboratory of Urban Spatial Information Engineering, ;;Hunan Provincial Key Laboratory of Geo-Information Engineering in Surveying;;Beijing Key Laboratory of Urban Spatial Information Engineering, Wuhan;;Xiangtan;;, ;;;;Beijing 430100;;411201;;100045  
Institution:  
长江大学地球科学学院  
Institution:  
长江大学地球科学学院;;自然资源部环鄱阳湖区域矿山环境监测与治理重点实验室  
Institution:  
长江大学地球科学学院;;湖南科技大学;;城市空间信息工程北京市重点实验室  
Institution:  
School of Geosciences, Yangtze University  
Institution:  
School of Geosciences, Yangtze University;;Key Laboratory of Mine Environmental Monitoring and Improving around Poyang Lake, Ministry of Natural Resources  
Institution:  
School of Geosciences, Yangtze University;;Mapping and Remote Sensing, Hunan University of Science and Technology;;Beijing Key Laboratory of Urban Spatial Information Engineering

### 153. Multi-Dimensional Fusion Deep Learning for Side Channel Analysis

摘要: The rapid advancement of deep learning has significantly heightened the threats posed by Side-Channel Attacks (SCAs) to information security, transforming their effectiveness to a degree several orders of magnitude superior to conventional signal processing techniques. However, the majority of existing Deep-Learning Side-Channel Attacks (DLSCAs) primarily focus on the classification accuracy of the trained model at the attack stage, often assuming that adversaries have unlimited computational and time resources during the profiling stage. This might result in an inflated assessment of the trained model's fitting capability in a real attack scenario. In this paper, we present a novel DLSCA model, called a Multi-Dimensional Fusion Convolutional Residual Dendrite (MD\_CResDD) network, to enhance and speed up the feature extraction process by incorporating a multi-scale feature fusion mechanism. By testing the proposed model on two software implementations of AES-128, we show that it is feasible to improve the profiling speed by at least 34% compared to other existing deep-learning models for DLSCAs and meanwhile achieved a certain level of improvement (8.4% and 0.8% for two implementations) in the attack accuracy. Furthermore, we also investigate how different fusion approaches, fusion times, and residual blocks can affect the attack efficiency on the same two datasets.  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
2079-9292  
入藏号:  
WOS:001116906000001  
Indexed Date:  
2023-12-17  
Institution Address:  
Hunan Univ Sci & Technol, Sch Phys & Elect Sci, Xiangtan 411201, Peoples R China  
Institution Address:  
Hunan Prov Key Lab Intelligent Sensors & Adv Sensi, Xiangtan 411201, Peoples R China  
Institution Address:  
Hunan Univ Sci & Technol, Sch Comp Sci & Engn, Xiangtan 411201, Peoples R China  
Institution:  
Institution:  
Hunan Prov Key Lab Intelligent Sensors & Adv Sensi  
Institution:

### 154. Underwater image object detection based on multi-scale feature fusion

摘要: Underwater object detection and classification technology is one of the most important ways for humans to explore the oceans. However, existing methods are still insufficient in terms of accuracy and speed, and have poor detection performance for small objects such as fish. In this paper, we propose a multi-scale aggregation enhanced (MAE-FPN) object detection method based on the feature pyramid network, including the multi-scale convolutional calibration module (MCCM) and the feature calibration distribution module (FCDM). First, we design the MCCM module, which can adaptively extract feature information from objects at different scales. Then, we built the FCDM structure to make the multi-scale information fusion more appropriate and to alleviate the problem of missing features from small objects. Finally, we construct the Fish Segmentation and Detection (FSD) dataset by fusing multiple data augmentation methods, which enriches the data resources for underwater object detection and solves the problem of limited training resources for deep learning. We conduct experiments on FSD and public datasets, and the results show that the proposed MAE-FPN network significantly improves the detection performance of underwater objects, especially small objects.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0932-8092  
eISSN:  
1432-1769  
入藏号:  
WOS:001303197600001  
Indexed Date:  
2024-09-07  
Institution Address:  
Southeast Univ, Sch Comp Sci & Engn, 2 Sipailou, Nanjing 210096, Jiangsu, Peoples R China  
Institution:

### 155. Invisible and robust watermarking model based on hierarchical residual

摘要: In current deep learning based watermarking technologies, it remains challenging to fully integrate the features of watermark and cover image. Most watermarking models with fixed-size kernel convolution exhibit restricted feature extraction ability, leading to incomplete feature fusion. To address this issue, a hierarchical residual fusion multi-scale convolution (HRFMS) module is designed. The method extracts image features from various receptive fields and implements feature interaction by residual connection. To produce watermarked image with high visual quality and attack resistance, a watermarking model based on the HRFMS is devised to achieve multi-scale feature fusion. Moreover, to minimize image distortion caused by watermark information, an attention mask layer is designed to guide the distribution of watermark information. The experimental results demonstrate that the invisibility and the robustness of the HRFMSNet are excellent. The watermarked images generated by the HRFMSNet are nearly visually indistinguishable from the cover images. The average peak signal-to-noise ratio of the watermarked images is 37.13 dB, and most of the bit error rates of the decoded messages are below 0.02.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:001356954100001  
Indexed Date:  
2024-11-23  
Institution Address:  
Nanchang Univ, Dept Elect Informat Engn, Nanchang 330031, Peoples R China  
Institution Address:  
Shanghai Univ Engn Sci, Sch Elect & Elect Engn, Shanghai 201620, Peoples R China  
Institution:  
Institution:

### 156. High-Resolution Image Building Extraction Based on Multi-level Feature

摘要: Objectives: The scale of buildings and their distribution is key indicators to measure the economic and social development of a region. Therefore, it is significant to study the extraction of buildings based on remote sensing images. Existing neural network methods still have shortcomings in the completeness of building extraction and the accuracy of building edges. To solve the above problems, this paper proposes a multi-level feature fusion network (MFFNet) based on high-resolution images. Methods: Firstly, we use edge detection operators to improve the ability of the network to recognize the boundaries of buildings. Secondly, we use a multi-path convolution fusion module to extract building features from multiple dimensions, and introduce a large receptive field convolution module to break through feature extraction. The process is limited by the size of the receptive field. After fusing the extracted features, the convolutional attention module is used to compress them, and the global features are further mined by pyramid pooling, so as to achieve high-precision extraction of buildings. Results: The current mainstream UNet, pyramid scene parsing network (PSPNet),multi attending path neural network (MAPNet) and multiscale-feature fusion deep neural networks with dilated convolution (MDNNet)are used as the comparison methods, and we use Wuhan University Aerial Image Dataset, Satellite Dataset II (East Asia) and Inria Aerial Image Dataset as experimental data for testing. Compared with the other four methods, MFFNet improves intersection over union, precision, recall, F1-score and mean average precision by 1.53%, 2.65%, 2.41%, 3.32% and 1.19% on average, achieves a better effect.Conclusions: MFFNet not only accurately captures the detail features of buildings, but also strengthens the extraction and utilization of global features. It has better extraction effect on large buildings and buildings in complex environment.  
摘要:  
建筑物规模及其分布是衡量一个地区经济社会发展状况的关键指标,因此研究基于遥感影像的建筑物提取具有重要意义。现有神经网络方法在建筑物提取的完整度、边缘精确度等方面仍存在不足,由此提出一种基于高分遥感影像的多层次特征融合网络(multi-level feature fusion network,MFFNet)。首先,利用边缘检测算子提升网络对建筑物边界的识别能力,同时借助多路径卷积融合模块多个维度提取建筑物特征,并引入大感受野卷积模块解决感受野大小对特征提取的限制问题;然后,对提取的特征进行融合,利用卷积注意力模块进行压缩,经金字塔池化进一步挖掘全局特征,从而实现建筑物的高精度提取。并与当前主流的UNet、 PSPNet(pyramid scene parsing network)、多路径特征融合网络(multi attending path neural network, MAPNet)和MDNNet(multiscale-feature fusion deep neural networks with dilated convolution)方法进行对比,使用亚米级的武汉大学航空影像数据集、卫星数据集II(东亚)与Inria航空影像数据集作为实验数据进行测试,结果发现,所提方法提取出的建筑物更为完整,边界更加精确。  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
2  
ISSN:  
1671-8860  
入藏号:  
CSCD:7309442  
Indexed Date:  
2023-01-15  
Institution Address:  
武汉大学遥感信息工程学院, 武汉, 湖北 430079, 中国  
Institution Address:  
中交第二公路勘察设计研究院有限公司, 武汉, 湖北 430056, 中国  
Institution Address:  
西南技术物理研究所, 成都, 四川 610041, 中国  
Institution Address:  
School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei 430079, China  
Institution Address:  
CCCC Second Highway Consultants Co. Ltd, Wuhan, Hubei 430056, China  
Institution Address:  
Norla Institute of Technical Physics, Chengdu, Sichuan 610041, China  
Institution:  
武汉大学遥感信息工程学院  
Institution:  
中交第二公路勘察设计研究院有限公司  
Institution:  
西南技术物理研究所  
Institution:  
School of Remote Sensing and Information Engineering, Wuhan University  
Institution:  
CCCC Second Highway Consultants Co. Ltd  
Institution:  
Norla Institute of Technical Physics

### 157. Multiscale progressive text prompt network for medical image

摘要: The accurate segmentation of medical images is a crucial step in obtaining reliable morphological statistics. However, training a deep neural network for this task requires a large amount of labeled data to ensure high-accuracy results. To address this issue, we propose using progressive text prompts as prior knowledge to guide the segmentation process. Our model consists of two stages. In the first stage, we perform contrastive learning on natural images to pretrain a powerful prior prompt encoder (PPE). This PPE leverages text prior prompts to generate multimodality features. In the second stage, medical image and text prior prompts are sent into the PPE inherited from the first stage to achieve the downstream medical image segmentation task. A multiscale feature fusion block (MSFF) combines the features from the PPE to produce multiscale multimodality features. These two progressive features not only bridge the semantic gap but also improve prediction accuracy. Finally, an UpAttention block refines the predicted results by merging the image and text features. This design provides a simple and accurate way to leverage multiscale progressive text prior prompts for medical image segmentation. Compared with using only images, our model achieves high-quality results with low data annotation costs. Moreover, our model not only has excellent reliability and validity on medical images but also performs well on natural images. The experimental results on different image datasets demonstrate that our model is effective and robust for image segmentation.& COPY; 2023 Elsevier Ltd. All rights reserved.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
7  
ISSN:  
0097-8493  
eISSN:  
1873-7684  
入藏号:  
WOS:001074358200001  
Indexed Date:  
2023-10-08  
Institution Address:  
Anhui Univ, Sch Comp Sci & Technol, Hefei, Peoples R China  
Institution Address:  
Sichuan Univ, Coll Comp Sci, Chengdu, Peoples R China  
Institution:  
Institution:

### 158. MDAR: A Multiscale Features-Based Network for Remotely Measuring Human

摘要: Remote photoplethysmography (rPPG) refers to a non-contact technique that measures heart rate through analyzing the subtle signal changes of facial blood flow captured by video sensors. It is widely used in contactless medical monitoring, remote health management, and activity monitoring, providing a more convenient and non-invasive way to monitor heart health. However, factors such as ambient light variations, facial movements, and differences in light absorption and reflection pose challenges to deep learning-based methods. To solve these difficulties, we put forward a measurement network of heart rate based on multiscale features. In this study, we designed and implemented a dual-branch signal processing framework that combines static and dynamic features, proposing a novel and efficient method for feature fusion, enhancing the robustness and reliability of the signal. Furthermore, we proposed an alternate time-shift module to enhance the model's temporal depth. To integrate the features extracted at different scales, we utilized a multiscale feature fusion method, enabling the model to accurately capture subtle changes in blood flow. We conducted cross-validation on three public datasets: UBFC-rPPG, PURE, and MMPD. The results demonstrate that MDAR not only ensures fast inference speed but also significantly improves performance. The two main indicators, MAE and MAPE, achieved improvements of at least 30.6% and 30.2%, respectively, surpassing state-of-the-art methods. These conclusions highlight the potential advantages of MDAR for practical applications.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001350991800001  
PubMed ID:  
39517688  
Indexed Date:  
2024-11-16  
Institution Address:  
Taiyuan Inst Technol, Dept Comp Engn, Taiyuan 030008, Peoples R China  
Institution Address:  
Taiyuan Normal Univ, Sch Comp Sci & Technol, Jinzhong 030619, Peoples R China  
Institution Address:  
Robert Gordon Univ, Sch Comp Engn & Technol, Aberdeen AB10 7QB, Scotland  
Institution Address:  
Beijing Jiaotong Univ, Sch Comp & Informat Technol, Beijing 100044, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 159. C2F-SemiCD: A Coarse-to-Fine Semi-Supervised Change Detection Method

摘要: A high-precision feature extraction model is crucial for change detection (CD). In the past, many deep learning-based supervised CD methods learned to recognize change feature patterns from a large number of labeled bi-temporal images, whereas labeling bi-temporal remote sensing images is very expensive and often time-consuming; therefore, we propose a coarse-to-fine semi-supervised CD method based on consistency regularization (C2F-SemiCD), which includes a coarse-to-fine CD network with a multiscale attention mechanism (C2FNet) and a semi-supervised update method. Among them, the C2FNet network "gradually" completes the extraction of change features from coarse-grained to fine-grained through multiscale feature fusion, channel attention mechanism, spatial attention mechanism, global context module, feature refine module, initial aggregation module, and final aggregation module. The semi-supervised update method uses the mean teacher method. The parameters of the student model are updated to the parameters of the teacher Model by using the exponential moving average (EMA) method. Through extensive experiments on three datasets and meticulous ablation studies, including crossover experiments across datasets, we verify the significant effectiveness and efficiency of the proposed C2F-SemiCD method.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
被引频次合计:  
11  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001184968700020  
Indexed Date:  
2024-04-18  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & Re, Wuhan 430079, Peoples R China  
Institution Address:  
Univ Tokyo, Grad Sch Frontier Sci, Chiba 2778561, Japan  
Institution:  
Institution:

### 160. MMNet: A multi-scale deep learning network for the left ventricular

摘要: With the development of deep learning network models, the automatic segmentation of medical images is becoming increasingly popular. Left ventricular cavity segmentation is an important step in the diagnosis of cardiac disease, but post-processing segmentation is a time-consuming and challenging task. That is why a fully automated segmentation method can assist specialists in increasing their efficiency. Inspired by the power of deep neural networks, a multi-scale multi-skip connection network (MMNet) model is proposed to fully automate the left ventricular segmentation of cardiac magnetic resonance imaging (MRI) images; this model is simple and efficient and has high segmentation accuracy without pre-detecting left ventricular localization. MMNet redesigns the classic encoder and decoder to take advantage of multi-scale feature information, effectively solving the problem of difficult segmentation due to blurred left ventricular edge information and the low accuracy of end-systolic segmentation of the cardiac area. In the model encoding stage, a multi-scale feature fusion module applying dilated convolution is proposed to obtain richer semantic information from different perceptual fields. The decoding stage reconstructs the full-size skip connection structure to make full use of the feature information obtained from different layers for contextual semantic information fusion. At the same time, a pre-activation module is used before each weighting layer to prevent overfitting phenomena from arising. The experimental results demonstrate that the proposed model has better segmentation performance than advanced benchmark models. Ablation experiments show that the proposed modules are effective at improving segmentation results. Therefore, MMNet is a promising approach for the left ventricular fully automated segmentation.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
30  
被引频次合计:  
33  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000682464100003  
Indexed Date:  
2021-08-06  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao 266590, Shandong, Peoples R China  
Institution Address:  
Shandong Univ Sci & Technol, Shandong Prov Key Lab Wisdom Min Informat Technol, Qingdao 266590, Shandong, Peoples R China  
Institution:  
Institution:

### 161. Bilevel Fusion With Local and Global Cues for Point Cloud Upsampling

摘要: This study focuses on point cloud upsampling, crucial in 3-D data processing but hindered by current 3-D sensor limitations. Point clouds from RGB-D cameras and light detection and ranging (LiDAR) scanners are often sparse, noisy, and irregular, challenging traditional processing methods reliant on prior knowledge and hindering detail preservation. Despite deep learning's transformative impact, issues like hole overfitting and insufficient local-global feature fusion persist. To address these, we introduce the bilevel fusion point cloud upsampling (BiPU) network. It features a parallel extractor for simultaneous local and global feature extraction and a consistency-based feature alignment module employing cross-attention for enhanced multiscale feature transfer. BiPU also incorporates 4-D encoding for rotational invariance and depthwise separable convolutions to reduce complexity and parameters. Tested across multiple datasets, BiPU excels in maintaining hole contours and reducing costs, marking a notable advancement in point cloud processing.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1551-3203  
eISSN:  
1941-0050  
入藏号:  
WOS:001303404700001  
Indexed Date:  
2024-09-07  
Institution Address:  
Tianjin Univ Technol, Sch Comp Sci & Engn, Tianjin 300384, Peoples R China  
Institution:

### 162. Multi-scale Feature Fusion Method for Spinal X-Ray Image Segmentation

摘要: In order to segment spine accurately from X-ray images,a method of spine X-ray image segmentation based on deep learning is proposed,which uses U-Net network based on multi-scale feature fusion.The convolutional layer in the U-Net model is replaced with an Inception network to extract features of different scales and perform multi-scale fusion.At the same time,the residual connection layer is added in front of the skip connection,and the convolution block attention module is added in front of the first up-sampling layer.In this paper,20 spinal X-ray images are verified by the model,and the Dice coefficient is 0.8457,which is 0.1351 higher than the recent method of spinal X-ray image segmentation.  
摘要:  
为了精确地从X线图像中分割脊柱,提出了一种基于深度学习的脊柱X线图像分割方法,使用基于多尺度特征融合的U-Net网络进行分割。将U-Net模型中的卷积层替换成类Inception网络来提取不同尺度的特征,并进行多尺度融合。同时在跳跃连接前增加残差连接层,并在首次上采样前添加卷积块注意力模块。该模型对20幅脊柱X线图像进行验证,Dice系数为0.845 7,与近期X线脊柱图像分割方法相比,提高了0.135 1。  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
被引频次合计:  
3  
ISSN:  
1002-8331  
入藏号:  
CSCD:6952052  
Indexed Date:  
2021-07-15  
Institution Address:  
云南大学信息学院, 昆明, 云南 650500, 中国  
Institution Address:  
School of Information,Yunnan University, Kunming, Yunnan 650500, China  
Institution:  
云南大学信息学院  
Institution:  
School of Information,Yunnan University

### 163. A Double Stream Person Re-Identification Method Based on Attention

摘要: Person re-identification mainly plays a role in multiple non-overlapping camera monitoring environments to determine whether the target person of interest that has appeared under a camera appears again under others. However, the image data, in the real scene, taken by the surveillance camera may be occluded or blurred, which increased the difficulty of identifying the pedestrian posturn and then lead to dramatically decrease the accuracy of recognition. To solve the above problems, we propose a dual branch multi-scale feature fusion network, which improves the expression ability of pedestrian features under partial occlusion by learning discriminative pedestrian features. By embedding the lightweight attention module into Residual neural network-50 (Resnet-50), the image sequence features of the channel dimension will be extracted, and the interference caused by the cluttered back ground information will be suppressed. In the training phase, the average pooling layer and the maximum pooling layer of different kernels and strides are utilized for different residual stages, and the mixed pooling strategy and mixed loss function of different kernels are designed. By comparison with the existing representative methods on Market1501, DukeMTMC reID and MTMS17 datasets, the experimental results show that the features extracted by the proposed method are more discriminative and with high recognition accuracy.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
ISSN:  
2169-3536  
入藏号:  
WOS:000936234100001  
Indexed Date:  
2023-03-12  
Institution Address:  
Chongqing Univ Sci & Technol, Sch Intelligent Technol & Engn, Chongqing 401331, Peoples R China  
Institution:

### 164. Ship Detection in SAR Images via Cross-Attention Mechanism

摘要: Deep learning has been widely applied to ship detection in Synthetic Aperture Radar (SAR) images. Unlike optical images, the current object detection methods have the problem of weak feature representation due to the low object resolution in SAR images. In addition, disturbed by chaotic noise, the features of classification and location are prone to significant differences, resulting in classification and location task misalignment. Therefore, this paper proposes a novel SAR ship target detection algorithm based on Cross-Attention Mechanism (CAM), which can establish the information interaction between the classification and localization task and strengthen the correlation between features through attention. In addition, to suppress the noise in multi-scale feature fusion, we designed an Attention-based Feature Fusion Module (AFFM), which uses the attention information between channels to perform the re-weighting operation. This operation can enhance useful feature information and suppress noise information. Experimental results show that on a benchmark SAR Ship Detection Dataset (SSDD), the Fully Convolutional One-Stage Object Detector (FCOS) with ResNet-50 backbone network was optimized to improve AP by 6.5% and computational cost by 0.51%. RetinaNet with ResNet-50 backbone network was optimized to improve AP by 1.8% and computational cost by 0.51%.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0703-8992  
eISSN:  
1712-7971  
入藏号:  
WOS:000854244000001  
Indexed Date:  
2022-09-21  
Institution Address:  
Xian Inst High Technol, Xian 710025, Peoples R China  
Institution:

### 165. Semantic Segmentation Network Based on Adaptive Attention and Deep

摘要: Deep learning has recently made significant progress in semantic segmentation. However, the current methods face critical challenges. The segmentation process often lacks sufficient contextual information and attention mechanisms, low-level features lack semantic richness, and high-level features suffer from poor resolution. These limitations reduce the model's ability to accurately understand and process scene details, particularly in complex scenarios, leading to segmentation outputs that may have inaccuracies in boundary delineation, misclassification of regions, and poor handling of small or overlapping objects. To address these challenges, this paper proposes a Semantic Segmentation Network Based on Adaptive Attention and Deep Fusion with the Multi-Scale Dilated Convolutional Pyramid (SDAMNet). Specifically, the Dilated Convolutional Atrous Spatial Pyramid Pooling (DCASPP) module is developed to enhance contextual information in semantic segmentation. Additionally, a Semantic Channel Space Details Module (SCSDM) is devised to improve the extraction of significant features through multi-scale feature fusion and adaptive feature selection, enhancing the model's perceptual capability for key regions and optimizing semantic understanding and segmentation performance. Furthermore, a Semantic Features Fusion Module (SFFM) is constructed to address the semantic deficiency in low-level features and the low resolution in high-level features. The effectiveness of SDAMNet is demonstrated on two datasets, revealing significant improvements in Mean Intersection over Union (MIOU) by 2.89% and 2.13%, respectively, compared to the Deeplabv3+ network.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001305075400001  
PubMed ID:  
39205000  
Indexed Date:  
2024-09-09  
Institution Address:  
Henan Polytech Univ, Sch Software, Jiaozuo 454000, Peoples R China  
Institution:

### 166. AMFANet: Advanced Deep Learning Techniques for High-Quality Image Style

摘要: This paper presents AMFANet, an advanced deep learning model engineered for high-quality image style transfer. AMFANet integrates cutting-edge techniques such as the Adaptive Multi-Scale Feature Fusion (AMSF) module and Hybrid Attention Mechanism (HAM) to significantly improve style consistency, content fidelity, and texture preservation. The model also utilizes Segmented Atrous Spatial Pyramid Pooling (SASPP) for effective multi-scale feature extraction. Comprehensive experimental evaluations demonstrate that AMFANet surpasses current state-of-the-art models like StyleGAN3, ChipGAN, ACL-GAN, and CycleGAN in generating high-fidelity stylized images while preserving intricate details and artistic essence. Future research will focus on optimizing computational efficiency, enabling multi-style transfer, enhancing user interaction, and exploring cross-domain applications. These findings highlight AMFANet's potential as a robust solution for advanced image style transfer in both artistic and practical domains.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0218-0014  
eISSN:  
1793-6381  
入藏号:  
WOS:001369614900001  
Indexed Date:  
2024-12-16  
Institution Address:  
Beijing Inst Graphic Commun, Sch New Media, 1 Sect 2,Xinghua St, Beijing 102600, Peoples R China  
Institution:  
Beijing Inst Graphic Commun

### 167. Defect Detection for Metal Shaft Surfaces Based on an Improved YOLOv5

摘要: To address the problem of low efficiency for manual detection in the defect detection field for metal shafts, we propose a deep learning defect detection method based on the improved YOLOv5 algorithm. First, we add a Convolutional Block Attention Module (CBAM) mechanism layer to the last layer of the backbone network to improve the feature extraction capability. Second, the neck network introduces the Bi-directional Feature Pyramid Network (BiFPN) module to replace the original Path-Aggregation Network (PAN) structure and enhance the multi-scale feature fusion. Finally, we use transfer learning to pre-train the model and improve the generalization ability of the model. The experimental results show that the method achieves an average accuracy of 93.6% mAP and a detection speed of 16.7 FPS for defect detection on the dataset, which can identify metal shaft surface defects quickly and accurately, and is of reference significance for practical industrial applications.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
eISSN:  
1424-8220  
入藏号:  
WOS:000968935300001  
PubMed ID:  
37050821  
Indexed Date:  
2023-05-02  
Institution Address:  
Wuhan Univ Sci & Technol, Key Lab Minist Educ Met Equipment & Control, Wuhan 430081, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Hubei Prov Key Lab Mech Transmiss & Mfg Engn, Wuhan 430081, Peoples R China  
Institution:  
Institution:

### 168. Building Multi-Feature Fusion Refined Network for Building Extraction

摘要: Deep learning approaches have been widely used in building automatic extraction tasks and have made great progress in recent years. However, the missing detection and wrong detection causing by spectrum confusion is still a great challenge. The existing fully convolutional networks (FCNs) cannot effectively distinguish whether the feature differences are from one building or the building and its adjacent non-building objects. In order to overcome the limitations, a building multi-feature fusion refined network (BMFR-Net) was presented in this paper to extract buildings accurately and completely. BMFR-Net is based on an encoding and decoding structure, mainly consisting of two parts: the continuous atrous convolution pyramid (CACP) module and the multiscale output fusion constraint (MOFC) structure. The CACP module is positioned at the end of the contracting path and it effectively minimizes the loss of effective information in multiscale feature extraction and fusion by using parallel continuous small-scale atrous convolution. To improve the ability to aggregate semantic information from the context, the MOFC structure performs predictive output at each stage of the expanding path and integrates the results into the network. Furthermore, the multilevel joint weighted loss function effectively updates parameters well away from the output layer, enhancing the learning capacity of the network for low-level abstract features. The experimental results demonstrate that the proposed BMFR-Net outperforms the other five state-of-the-art approaches in both visual interpretation and quantitative evaluation.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
28  
被引频次合计:  
31  
eISSN:  
2072-4292  
入藏号:  
WOS:000677104700001  
Indexed Date:  
2021-08-01  
Institution Address:  
Yangtze Univ, Sch Geosci, Wuhan 430100, Peoples R China  
Institution Address:  
Beijing Key Lab Urban Spatial Informat Engn, Beijing 100045, Peoples R China  
Institution Address:  
Hunan Univ Sci & Technol, Hunan Prov Key Lab Geoinformat Engn Surveying Map, Xiangtan 411201, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Key Lab Earth Observat Hainan Prov, Sanya 572029, Peoples R China  
Institution:  
Institution:  
Beijing Key Lab Urban Spatial Informat Engn  
Institution:  
Institution:  
Institution:  
Key Lab Earth Observat Hainan Prov

### 169. Detection Method for Rice Seedling Planting Conditions Based on Image

摘要: In response to the need for precision and intelligence in the assessment of transplanting machine operation quality, this study addresses challenges such as low accuracy and efficiency associated with manual observation and random field sampling for the evaluation of rice seedling planting conditions. Therefore, in order to build a seedling insertion condition detection system, this study proposes an approach based on the combination of image processing and deep learning. The image processing stage is primarily applied to seedling absence detection, utilizing the centroid detection method to obtain precise coordinates of missing seedlings with an accuracy of 93.7%. In the target recognition stage, an improved YOLOv8 Nano network model is introduced, leveraging deep learning algorithms to detect qualified and misplaced seedlings. This model incorporates ASPP (atrous spatial pyramid pooling) to enhance the network's multiscale feature extraction capabilities, integrates SimAM (Simple, Parameter-free Attention Module) to improve the model's ability to extract detailed seedling features, and introduces AFPN (Asymptotic Feature Pyramid Network) to facilitate direct interaction between non-adjacent hierarchical levels, thereby enhancing feature fusion efficiency. Experimental results demonstrate that the enhanced YOLOv8n model achieves precision (P), recall (R), and mean average precision (mAP) of 95.5%, 92.7%, and 95.2%, respectively. Compared to the original YOLOv8n model, the enhanced model shows improvements of 3.6%, 0.9%, and 1.7% in P, R, and mAP, respectively. This research provides data support for the efficiency and quality of transplanting machine operations, contributing to the further development and application of unmanned field management in subsequent rice seedling cultivation.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
eISSN:  
2076-3417  
入藏号:  
WOS:001191740900001  
Indexed Date:  
2024-04-03  
Institution Address:  
Qingdao Agr Univ, Coll Mech & Elect Engn, Qingdao 266109, Peoples R China  
Institution Address:  
Natl Key Lab Agr Equipment Technol, Beijing 100083, Peoples R China  
Institution:  
Institution:  
Natl Key Lab Agr Equipment Technol

### 170. DCSN: Deep Compressed Sensing Network for Efficient Hyperspectral Data

摘要: Requirements of compressed sensing techniques targeted at miniaturized hyperspectral satellite applications include lightweight onboard hardware, high-speed sensing, low sampling rate for compressing the massive volume of typical hyperspectral data, and noise robustness for reliable data transmission to the ground station. We achieve all these aims via deep learning, and neural networks resulted from which can be implemented on-chip, thereby allowing light hardware implementation. Our neural networks were trained from small-scaled data, but, even so, the resulting encoder achieves a very low sampling rate and very high speed. Unlike typical network training, the input-output pairs are not square but stripe-like images, partly because compressed acquisition does not allow performing compression after obtaining complete data cube and partly because stripe-like acquisition well matches the popular pushbroom hyperspectral sensing schemes. Even with such hard restriction caused by nontraditional training, the resulting decoder still reconstructs the image with high accuracy. To match the requirement of pushbroom sensing, a lightweight encoder is proposed to compress the stripe-like images immediately. Meanwhile, multiscale feature fusion block (MFB) and aggregation (MFA) modules are proposed to form our decoder for enhancing the feature representation of the compressed acquisitions. Furthermore, we achieve joint spatial/spectral super-resolution (SR) progressively, ensuring accurate hyperspectral reconstruction via a low-rank-driven decoder. The encoder and decoder are trained in an end-to-end manner, where noise robustness is forced during the training stage. Comprehensive experiments demonstrate the superiority of the proposed hyperspectral compressed sensing method, as well as its one-shot transfer learning (OTL)-based extension, both quantitatively and qualitatively.  
Web of Science 核心合集中的 "被引频次":  
31  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
32  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000690968800050  
Indexed Date:  
2021-09-10  
Institution Address:  
Natl Pingtung Univ Sci & Technol, Dept Management Informat Syst, Pingtung 91201, Taiwan  
Institution Address:  
Natl Cheng Kung Univ, Dept Elect Engn, Tainan 70101, Taiwan  
Institution:  
Institution:

### 171. AVA-YOLO: image-based multiscale feature fusion enhanced perception

摘要: The global climate change has led to frequent occurrences of snow avalanche disasters. However, the significant variations in scale and shape during the avalanche process, and complex background imagery pose significant challenges to automated detection efforts. There is an urgent need to combine advanced deep learning technology to research automatic detection and recognition of avalanches in the field. In this paper, a novel deep learning model based on YOLOv8 improved multi-scale detection called AVA-YOLO is proposed to solve this problem. In AVA-YOLO, a key component, AKA (AKConv Combined Attention) module was designed and developed. This module combines the deformable convolutional properties of AKConv with the state-of-the-art self-attention module Exponential Moving Average, aiming to better perceive the feature map information of different shaped avalanches and to enhance the global relevance, thus improving the utilization of the information. Secondly, a new multi-scale sensing network structure was designed by increasing the number of detection heads to four and introducing the AKA module into the key positions of the network, while the association between model layers was newly designed to enhance the fusion of shallow and deep information to improve the detection accuracy. Experimental results demonstrated the effectiveness of AVA-YOLO, achieving 95.7% mAP50 and 75.6% mAP50:95 detection accuracies, as well as an F1 score of 0.92. Finally, a number of experiments were conducted to demonstrate the superior performance of the proposed model in comparison to other versions of YOLO, which will further exploit the potential of webcams as an underutilized technical capability in snow avalanche intelligence and portable monitoring.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001324012500001  
Indexed Date:  
2024-10-07  
Institution Address:  
Chengdu Univ Technol, State Key Lab Geohazard Prevent & Geoenvironm Pro, Chengdu 610059, Peoples R China  
Institution Address:  
Chengdu Univ Technol, Coll Nucl Technol & Automat Engn, Chengdu 610059, Peoples R China  
Institution Address:  
Chengdu Univ Technol, Coll Comp & Cyber Secur, Chengdu 610059, Peoples R China  
Institution Address:  
Chengdu Univ Technol, Sichuan Engn Technol Res Ctr Ind Internet Intelli, Chengdu 610059, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 172. Geological remote sensing interpretation via a local-to-global sensitive

摘要: Interpreting surface geological elements (such as rocks, minerals, soils, and water bodies) is the main task of geological survey, which plays a crucial role in geological environment remote sensing (GERS). However, the characteristics of geological elements, including high variabilities, various morphology, complicated boundaries and imbalanced class distribution, make it still a challenge for deep learning methods to interpret GERS images. Considering the correlations of geological elements as the regionalized variables in geostatistics, the sensitive features of GERS interpretation mainly include three aspects: tonal, textural and structural characteristics within a singular-class elements, spatial and spectral correlations of adjacent elements, and their global tectonic or spatial distribution. Thus, to simulate the manual interpretation process of geologists from local to global and promote GERS interpretation performance, we propose a local-to-global multi-scale feature fusion network (LGMSFNet). A geological object context represents the intra-class semantic dependencies of pixel sets with the same class. And a local feature aggregation module models the channel and spatial association. Then discriminative features are integrated by a global feature fusion module. For the model optimization, we focus on hard examples during the training process to achieve the balanced optimization of various categories. Two research areas that include large-scale rocks, soils and water exposed on the surface are selected. Massive experiments demonstrate the superiority of the LGMSFNet in GERS interpretation.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1569-8432  
eISSN:  
1872-826X  
入藏号:  
WOS:001363526300001  
Indexed Date:  
2024-12-01  
Institution Address:  
China Univ Geosci, Sch Comp Sci, Wuhan 430078, Peoples R China  
Institution:

### 173. MDANet: A High-Resolution City Change Detection Network Based on

摘要: In the domains of geographic information systems and remote sensing image analysis, change detection is vital for examining surface variations in high-resolution remote sensing pictures. However, the intricate texture characteristics and rich details found in high-resolution remote sensing photos are difficult for conventional change detection systems to deal with. Target misdetection, missed detections, and edge blurring are further problems with current deep learning-based methods. This research proposes a high-resolution city change detection network based on difference and attention mechanisms under multi-scale feature fusion (MDANet) to address these issues and improve the accuracy of change detection. First, to extract features from dual-temporal remote sensing pictures, we use the Siamese architecture as the encoder network. The Difference Feature Module (DFM) is employed to learn the difference information between the dual-temporal remote sensing images. Second, the extracted difference features are optimized with the Attention Refinement Module (ARM). The Cross-Scale Fusion Module (CSFM) combines and enhances the optimized attention features, effectively capturing subtle differences in remote sensing images and learning the finer details of change targets. Finally, thorough tests on the BTCDD dataset, LEVIR-CD dataset, and CDD dataset show that the MDANet algorithm performs at a cutting-edge level.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
11  
eISSN:  
2072-4292  
入藏号:  
WOS:001258371800001  
Indexed Date:  
2024-07-04  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Management Sci & Engn, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Forestry Univ, Coll Informat Sci & Technol, Nanjing 210037, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Jiangsu Key Lab Big Data Anal Technol, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Collaborat Innovat Ctr Atmospher Environm & Equipm, Nanjing 210044, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 174. Optimizing cardiovascular image segmentation through integrated

摘要: BACKGROUND: Cardiovascular diseases are the top cause of death in China. Manual segmentation of cardiovascular images, prone to errors, demands an automated, rapid, and precise solution for clinical diagnosis.  
OBJECTIVE: The paper highlights deep learning in automatic cardiovascular image segmentation, efficiently identifying pixel regions of interest for auxiliary diagnosis and research in cardiovascular diseases.  
METHODS: In our study, we introduce innovative Region Weighted Fusion (RWF) and Shape Feature Refinement (SFR) modules, utilizing polarized self-attention for significant performance improvement in multiscale feature integration and shape fine-tuning. The RWF module includes reshaping, weight computation, and feature fusion, enhancing high-resolution attention computation and reducing information loss. Model optimization through loss functions offers a more reliable solution for cardiovascular medical image processing.  
RESULTS: Our method excels in segmentation accuracy, emphasizing the vital role of the RWF module. It demonstrates outstanding performance in cardiovascular image segmentation, potentially raising clinical practice standards.  
CONCLUSIONS: Our method ensures reliable medical image processing, guiding cardiovascular segmentation for future advancements in practical healthcare and contributing scientifically to enhanced disease diagnosis and treatment.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0928-7329  
eISSN:  
1878-7401  
入藏号:  
WOS:001255718100035  
PubMed ID:  
38759064  
Indexed Date:  
2024-08-09  
Institution Address:  
Xiamen Univ, Sch Informat, Xiamen 36100, Fujian, Peoples R China  
Institution Address:  
Xiamen Univ, Xiamen Cardiovasc Hosp Xiamen Univ, Emergency Dept, Xiamen, Fujian, Peoples R China  
Institution Address:  
Changji Univ, Sch Informat Engn, Changji 831100, Xinjiang Uygur, Peoples R China  
Institution:  
Institution:  
Institution:

### 175. Fast-SegNet: fast semantic segmentation network for small objects

摘要: Semantic segmentation is a fundamental step in image understanding, playing a crucial role in the fields of automatic driving, medical image analysis, defect detection, etc. Despite significant progress in deep learning-based image segmentation, challenges in terms of accuracy and efficiency still exist, especially for small-scale objects. In this paper, we present a novel data augmentation method for small-scale objects in images, aiming to address the issue of class imbalance. Specifically, we extract small-scale objects from one image and then copy-scale-and-paste them to other images. Additionally, a novel multi-scale feature fusion module is proposed to effectively combine features from both deep and shallow neural network layers. Subsequently, the data augmentation method and multi-scale feature fusion module are utilized in the proposed Fast-SegNet architecture for semantic segmentation. Extensive experiments demonstrate that Fast-SegNet could improve segmentation performance, especially for small-scale objects with an acceptable computational cost. State-of-the-art performance has been achieved on CamVid, CityScapes, and MOST (Micro-optical sectioning tomography) datasets with respect to the tradeoff between accuracy and speed. Specifically, the CamVid dataset yields mean IoU (Intersection over Union) values of 45.7% and 38.6% for small-scale objects as Pedestrian and Bicyclist, respectively. The CityScapes dataset demonstrates mean IoU of 43.43% and 43.56% for small-scale objects as Traffic Light and Rider, respectively. The MOST dataset results in a segmentation mean IoU of 88.2% for vessels in the mouse brain. In conclusion, our approach achieves better results in terms of accuracy and efficiency on three datasets. Codes are available at https://github.com/apple1986/Fast-SegNet.  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001178219700002  
Indexed Date:  
2024-03-31  
Institution Address:  
Wuhan Inst Technol, Sch Comp Sci & Engn, Hubei Key Lab Intelligent Robot, Wuhan 430205, Hubei, Peoples R China  
Institution Address:  
Huazhong Univ Sci & Technol, Britton Chance Ctr Biomed Photon, Wuhan Natl Lab Optoelect, Wuhan 430074, Hubei, Peoples R China  
Institution:  
Institution:

### 176. Cross-site scripting detection with two-channel feature fusion emb e dde

摘要: In the era of big data, stealing users' private data has become one of the main targets of network hack-ers. In recent years, cross-site scripting (XSS) attacks to obtain users' privacy data have been one of the main web attack methods of network hackers. Traditional antivirus software cannot identify such cross-site scripting attacks. To identify cross-site scripting attacks quickly and accurately, we proposed a cross-site scripting detection model (C-BLA) with two-channel multi-scale feature fusion embedded in a self-attention mechanism. The model first maps cross-site scripting payloads into spatial vectors by data preprocessing using Word2Vec. Then the two-channel network performs feature extraction on the data. Channel I: extract local features of cross-site scripting payloads at different scales by designing par-allel one-dimensional convolutional layers with different convolutional kernel sizes; Channel II: extract semantic information of cross-site scripting payloads from two directions of positive and negative order using a bidirectional Long-Short Term Memory network, and then embed the self-attention mechanism to strengthen the semantic information features. Experiments show that the proposed model achieves a precision rate of 99.8 % and a recall rate of 99.1 % for cross-site scripting detection, which is a certain im-provement in detection rate compared with a single deep learning model and traditional machine learn-ing methods. The two-channel feature fusion of this model better solves the cross-site scripting detection problem.(c) 2022 Elsevier Ltd. All rights reserved.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
9  
被引频次合计:  
10  
ISSN:  
0167-4048  
eISSN:  
1872-6208  
入藏号:  
WOS:000901150600014  
Indexed Date:  
2023-01-16  
Institution Address:  
Shanghai Normal Univ, Coll Informat Mech & Elect Engn, Shanghai, Peoples R China  
Institution:

### 177. MMMNet: An End-to-End Multi-Task Deep Convolution Neural Network With

摘要: As the evaluation of image quality depends on the human visual system (HVS), many existing image quality assessment (IQA) methods focus on modeling the HVS to account for subjective perception. The visual attention of the HVS makes humans more sensitive to distortion on the attended regions than on regions which are not the focus of attention. Therefore, we propose an end-to-end multi-task deep convolution neural network with multi-scale and multi-hierarchy fusion (MMMNet), in which the IQA and saliency subtasks are jointly optimized to improve saliency-guided IQA performance. Particularly, the incorporation of saliency information is achieved by fusing saliency features with IQA features hierarchically to progressively improve the IQA features over network depth. A multi-scale feature extraction module (MSFE) is proposed to provide effective saliency features for the IQA network. Based on the saliency fusion, MMMNet introduces an auxiliary saliency task, achieving the multi-task learning to improve the generalization of the IQA task. Experimental results show that MMMNet achieves state-of-the-art performance and strong generalization ability on IQA databases.  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
34  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
35  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:000725812500023  
Indexed Date:  
2021-12-10  
Institution Address:  
Xi An Jiao Tong Univ, Sch Informat & Commun Engn, Minist Educ, Key Lab Intelligent Networks & Network Secur, Xian 710049, Peoples R China  
Institution Address:  
Univ Calif San Diego, Dept Elect & Comp Engn, La Jolla, CA 92093 USA  
Institution:  
Institution:

### 178. LMANet: A Lightweight Asymmetric Semantic Segmentation Network Based on

摘要: With the swift progress of deep learning and its wide application in semantic segmentation, the effect of semantic segmentation has been significantly improved. However, how to achieve a reasonable compromise between accuracy, model size, and inference speed is crucial. In this paper, we propose a lightweight multi-scale asymmetric encoder-decoder network (LMANet) that is designed on the basis of an encoder-decoder structure. First, an optimized bottleneck module is used to extract features from different levels, and different receptive fields are applied to obtain effective information on different scales. Then, a channel-attention module and a feature-extraction module are introduced to constitute the residual structure, and different feature maps are connected by a feature-fusion module to effectively improve segmentation accuracy. Finally, a lightweight multi-scale decoder is designed to recover the image, and a spatial attention module is added to recover the spatial details effectively. This paper has verified the proposed method on the Cityscapes dataset and CamVid dataset and achieved mean intersection over union (mIoU) of 73.9% and 71.3% with the inference speeds of 111 FPS and 118 FPS, respectively, and the number of parameters is only 0.85 M.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001311244700001  
Indexed Date:  
2024-09-19  
Institution Address:  
Anhui Univ Sci & Technol, Sch Comp Sci & Engn, Huainan 232001, Peoples R China  
Institution Address:  
China Telecom Co Ltd, Anhui Branch, Hefei 230000, Peoples R China  
Institution:  
Institution:

### 179. Efficient-Residual Net-A Hybrid Neural Network for Automated Brain Tumor

摘要: A multiscale feature fusion of Efficient-Residual Net is proposed for classifying tumors on brain Magnetic resonance images with solid or cystic masses, inadequate borders, unpredictable cystic and necrotic regions, and variable heterogeneity. Therefore, in this research, Efficient-Residual Net is proposed by efficaciously amalgamating features of two Deep Convolution Neural Networks-ResNet50 and EffficientNetB0. The skip connections in ResNet50 have reduced the chances of vanishing gradient and overfitting problems considerably thus learning of a higher number of features from input MR images. In addition, EffficientNetB0 uses a compound scaling coefficient for uniformly scaling the dimensions of the network such as depth, width, and resolution. The hybrid model has improved classification results on brain tumors with similar morphology and is tested on three internet repository datasets, namely, Kaggle, BraTS 2018, BraTS 2021, and real-time dataset from Post Graduate Institute of Medical Education and Research (PGIMER), Chandigarh. It is observed that the proposed system delivers an overall accuracy of 96.40%, 97.59%, 97.75%, and 97.99% on the four datasets, respectively. The proposed hybrid methodology has given assuring results of 98%-99% of other statistical such parameters as precision, negatively predicted values, and F1 score. The cloud-based web page is also created using the Django framework in Python programming language for accurate prediction and classification of different types of brain tumors.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0899-9457  
eISSN:  
1098-1098  
入藏号:  
WOS:001312062500001  
Indexed Date:  
2024-09-21  
Institution Address:  
TIET, Dept Elect & Instrumentat Engn, Patiala, India  
Institution Address:  
PGIMER, Dept Radiodiag & Imaging, Chandigarh, India  
Institution:  
Institution:

### 180. Vehicle-assisted bridge damage assessment by combining attention

摘要: Vehicle-assisted bridge damage identification has great application potential, but it is still difficult to extract damage-sensitive features from multi-source monitoring data and accurately evaluate the bridge damage status. To solve this problem, an Attention-LSTM-based Feature Fusion Model (ALFF-Net) is proposed. The model improves the perception ability of Bi-LSTM cells for multi-scale feature information in time series data through a preset data reconstruction layer. Furthermore, by employing attention mechanism and feature fusion strategy, the model reduces the prediction difficulty of downstream branches of deep neural networks and further improves the modeling ability for the important dependency relationships in the sequence data. A monitoring dataset under different road roughness and vehicle speeds is generated through a vehicle-bridge interaction system simulation, and the bridge damage identification performance of the ALFF-Net model is comprehensively tested. The results show that the ALFFNet model improves the damage identification accuracy by up to 19.30% compared to the classical LSTM network while significantly reducing computational costs, and the identification errors under different road roughness levels are less than 3%. Moreover, by comparing the identification accuracy of the ALFF-Net model under different data-driven schemes, the robustness of the bridge damage detection results with synergistic multi-source monitoring data is verified.  
摘要:  
基于车辆辅助的桥梁损伤识别具有巨大应用潜力,但仍难以从多源监测数据中提取损伤敏感特征,进而准确评估桥梁损伤状态。为此,提出了基于长短时记忆网络的注意力加权特征融合模型(ALFF-Net)。该模型通过预置数据重构层,提高了Bi-LSTM单元对时间序列多尺度特征信息的感知能力。同时结合注意力机制和特征融合策略,降低了深度神经网络下游分支的预测难度,进一步提升了模型对序列数据重要依赖关系的建模能力。通过车-桥耦合仿真生成了不同路面不平整度和车速下的监测数据集,对ALFF-Net模型的桥梁损伤识别性能进行综合测试。结果表明:ALFF-Net模型较经典LSTM网络在显著降低计算成本的同时,损伤识别准确率最高可提升19.30%,且各级路面不平整度下的识别误差均小于3%。进一步地,通过对比ALFF-Net模型在不同监测数据驱动方案下的识别精度,验证了协同多源监测数据的桥梁结构损伤检测结果更为鲁棒。  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1004-4523  
入藏号:  
CSCD:7798323  
Indexed Date:  
2024-11-08  
Institution Address:  
东南大学土木工程学院, 南京, 江苏 211189, 中国  
Institution Address:  
东南大学;;东南大学土木工程学院;;东南大学, 混凝土与预应力混凝土结构教育部重点实验室;;;;智慧建造与运维国家地方联合工程研究中心, 南京;;南京;;南京, 江苏;;江苏;;江苏 210096;;211189;;211189, 中国  
Institution Address:  
School of Civil Engineering, Southeast University, Nanjing, Jiangsu 211189, China  
Institution Address:  
Southeast University;;School of Civil Engineering, Southeast University;;Southeast University, Key Laboratory of Concrete and Prestressed Concrete Structures of Ministry of Education;;;;National and Local Joint Engineering Research Center for Intelligent Construction and Maintenance, Nanjing;;Nanjing;;Nanjing, ;;;; 210096;;211189;;211189  
Institution:  
东南大学土木工程学院  
Institution:  
东南大学;;东南大学土木工程学院;;东南大学  
Institution:  
School of Civil Engineering, Southeast University  
Institution:  
Southeast University;;School of Civil Engineering, Southeast University;;Southeast University

### 181. Multi-Scale Feature Fusion Attention Network Model for Cervical

摘要: In recent years,deep learning has been increasingly employed to effectively assist the diagnosis of medical images. However,common medical cone segmentation models cannot extract cervical spine details,resulting in incomplete cone segmentation or relatively blurred edges. To improve the segmentation accuracy of cervical vertebrae Magnetic Resonance Imaging(MRI),a Multi-Scale Feature Fusion Attention(MSFFA)network model based on ResNet is proposed. The multi-scale attention module fuses different receptive fields for attention feature enhancement. Further, to reduce the loss of feature information fusion,a cross-scale feature fusion module is used to enhance the features of similar and edge domains. Finally,the feature information of the original sample is integrated into the segmentation results for detail enhancement,further optimizing the segmentation performance of the model. Experimental results show that compared with U-Net,AttUNet,and other models,the MSFFA model provides a more complete cervical vertebrae structure and smoother edges and can achieve more accurate segmentation results in lumbar vertebrae segmentation. Compared with the best model,DeepLabv3+,the mean Dice Similarity Coefficient(DSC)increases by 1.05 percentage points.  
摘要:  
近年来,基于深度学习的医学图像辅助诊断逐渐成为主流,但常见的医疗锥体分割模型缺乏对颈椎细节信息的提取,导致锥体分割不完整或边缘相对模糊。为了提高颈椎MRI图像的分割精度,基于ResNet构建一种多尺度特征融合注意力(MSFFA)网络模型。利用多尺度注意力模块融合不同感受野进行注意力特征增强,同时为了降低特征信息融合的损耗,采用跨尺度特征融合模块进行相似域和边缘域特征增强,最终将原始样本的特征信息整合到分割结果中进行细节增强,进一步优化模型分割性能。实验结果表明,MSFFA模型相比于U-Net、AttUNet等模型分割得到的颈椎结构更完整、边缘更平滑,同时在腰椎分割中也能取得更精确的分割结果,并且相比于最优对照模型DeepLabv3+,Dice相似系数的均值提升了1.05个百分点。  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
1  
ISSN:  
1000-3428  
入藏号:  
CSCD:7573844  
Indexed Date:  
2024-01-06  
Institution Address:  
西南石油大学计算机科学学院, 成都, 四川 610500, 中国  
Institution Address:  
四川大学华西医院骨科, 成都, 四川 610041, 中国  
Institution Address:  
School of Computer Science,Southwest Petroleum University, Chengdu, Sichuan 610500, China  
Institution Address:  
Department of Orthopedics,West China Hospital of Sichuan University, Chengdu, Sichuan 610041, China  
Institution:  
西南石油大学计算机科学学院  
Institution:  
四川大学华西医院骨科  
Institution:  
School of Computer Science,Southwest Petroleum University  
Institution:  
Department of Orthopedics,West China Hospital of Sichuan University

### 182. Voids Filling of DEM with Multiattention Generative Adversarial Network

摘要: The digital elevation model (DEM) acquired through photogrammetry or LiDAR usually exposes voids due to phenomena such as instrumentation artifact, ground occlusion, etc. For this reason, this paper proposes a multiattention generative adversarial network model to fill the voids. In this model, a multiscale feature fusion generation network is proposed to initially fill the voids, and then a multiattention filling network is proposed to recover the detailed features of the terrain surrounding the void area, and the channel-spatial cropping attention mechanism module is proposed as an enhancement of the network. Spectral normalization is added to each convolution layer in the discriminator network. Finally, the training of the model by a combined loss function, including reconstruction loss and adversarial loss, is optimized. Three groups of experiments with four different types of terrains, hillsides, valleys, ridges and hills, are conducted for validation of the proposed model. The experimental results show that (1) the structural similarity surrounding terrestrial voids in the three types of terrains (i.e., hillside, valley, and ridge) can reach 80-90%, which implies that the DEM accuracy can be improved by at least 10% relative to the traditional interpolation methods (i.e., Kriging, IDW, and Spline), and can reach 57.4%, while other deep learning models (i.e., CE, GL and CR) only reach 43.2%, 17.1% and 11.4% in the hilly areas, respectively. Therefore, it can be concluded that the structural similarity surrounding the terrestrial voids filled using the model proposed in this paper can reach 60-90% upon the types of terrain, such as hillside, valley, ridge, and hill.  
在 BIOSIS Citation Index 中的被引频次:  
6  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
42  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
1  
被引频次合计:  
42  
eISSN:  
2072-4292  
入藏号:  
WOS:000768282100001  
Indexed Date:  
2022-03-24  
Institution Address:  
Guilin Univ Technol, Guangxi Key Lab Spatial Informat & Geomat, Guilin 541004, Peoples R China  
Institution Address:  
Guilin Univ Technol, Coll Geomat & Geoinformat, Guilin 541004, Peoples R China  
Institution Address:  
Guilin Univ Technol, Coll Earth Sci, Guilin 541004, Peoples R China  
Institution:  
Institution:  
Institution:

### 183. A Unified Cloud Detection Method for Suomi-NPP VIIRS Day and Night PAN

摘要: Cloud detection is a necessary step before the application of remote sensing images. However, the radiation intensity similarity between artificial lights and clouds is higher in nighttime remote sensing images than in daytime remote sensing images, making it difficult to distinguish artificial lights from clouds. This article proposes a deep learning method called multifeature fusion for cloud detection network (MFFCD-Net) to detect clouds in daytime and nighttime remote sensing images. A dilated residual upsampling module was designed for upsampling feature maps while enlarging the receptive field. A multiscale feature-extraction fusion module (MFEF) was designed to enhance the ability to distinguish regular textures of artificial lights from random textures of clouds. Moreover, an adaptive feature-fusion module (AFF) was designed to select and fuse the feature in the encoding stage and decoding stage, thus improving the cloud detection accuracy. To the best of our knowledge, this is the first time that a method is designed for cloud detection in both daytime and nighttime remote sensing images. The experimental results on Suomi-NPP Visible Infrared Imaging Radiometer Suite (VIIRS) of the panchromatic (PAN) day/night band (DNB) images show that MFFCD-Net could obtain a better balance in commission and omission rates than baseline methods (92.3% versus 90.5% on the F1-score) in daytime remote sensing images. Although artificial lights introduce strong interference in nighttime remote sensing images, MFFCD-Net can better distinguish artificial lights from clouds than baseline methods (90.8% versus 88.4% on the F1-score). The results indicate that MFFCD-Net is promising for cloud detection both in daytime and nighttime remote sensing images. The source code and dataset are available at https://github.com/Neooolee/MFFCD-Net.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001276870500034  
Indexed Date:  
2024-08-09  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Astronaut, Nanjing 210016, Peoples R China  
Institution Address:  
Army Engn Univ PLA, Inst Commun Engn, Nanjing 210001, Peoples R China  
Institution Address:  
Beijing Inst Control Engn, Beijing 100190, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Dept Land Surveying & Geoinformat, Hong Kong, Peoples R China  
Institution Address:  
VTT Tech Res Ctr Finland Ltd, Espoo 02044, Finland  
Institution:  
Institution:  
Institution:  
Beijing Inst Control Engn  
Institution:  
Institution:

### 184. Improved brachial plexus nerve segmentation method based on multi-scale

摘要: With the features of low Signal-to-Noise Ratio (SNR) and blurred edges,ultrasound images of the brachial plexus nerve are hard to be segmented manually.Although some results have been gained by existing segmentation models,the segmentation effect is not satisfied due to the small target area and irregular shape of the brachial plexus nerve structure.Aiming at the above problems,a multi-scale feature fusion-based brachial plexus nerve segmentation model was proposed,namely Nerve-segmentation Feature Pyramid Network (Ner-FPN).In the feature extraction stage,an Xception-like structure was designed for multi-scale feature extraction.In the prediction segmentation stage,a bidirectional FPN structure was used for feature fusion prediction.The BP (Brachial Plexus) dataset from the Kaggle brachial plexus nerve ultrasound image segmentation competition was used as the experimental data.The experimental results show that compared with the mainstream deep learning segmentation models U-Net and SegNet (Segmentation Network),the Dice Similar Coefficient(DSC) of Ner-FPN model for brachial plexus nerve segmentation can reach 0.703,which is 10.7 percentage points and 14.5 percentage points higher than those of U-Net and SegNet,and 5.5 percentage points and 3.4 percentage points higher than those of improved models QU-Net and Efficient+U-Net in the same dataset,verifying that the proposed model can be an aid for diagnosis.  
摘要:  
臂丛神经超声影像信噪比(SNR)低、边缘模糊且人工分割难度较大。现有的分割模型虽然取得了一些成果,但碍于臂丛神经结构目标区域小、形状不规则,分割效果欠佳。针对上述问题,设计基于多尺度特征融合的臂丛神经分割模型,即针对神经部位分割的特征金字塔网络(Ner-FPN)。在特征提取阶段,设计一种仿Xception的结构进行多尺度特征提取;在预测分割阶段,采用双向FPN结构进行特征融合预测。在Kaggle臂丛神经超声影像分割竞赛的BP数据集上的实验结果表明,Ner-FPN模型对臂丛神经分割的Dice相似系数(DSC)可达0.703,与主流的深度学习分割模型U-Net、SegNet相比,分别提高了10.7个百分点和14.5个百分点,对比相同数据集中的其他改进模型QU-Net和Efficient+U-Net,DSC分别提高了5.5个百分点和3.4个百分点,可见所提模型能够起到辅助诊断的效果。  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1001-9081  
入藏号:  
CSCD:7403600  
Indexed Date:  
2023-06-02  
Institution Address:  
青岛科技大学信息科学技术学院, 青岛, 山东 266061, 中国  
Institution Address:  
College of Information Science and Technology,Qingdao University of Science and Technology, Qingdao, Shandong 266061, China  
Institution:  
青岛科技大学信息科学技术学院  
Institution:  
College of Information Science and Technology,Qingdao University of Science and Technology

### 185. MSMP-Net: A Multi-Scale Neural Network for End-to-End Monkeypox Virus

摘要: Monkeypox is a zoonotic disease caused by monkeypox virus infection. It is easily transmitted among people and poses a major threat to human health, making it of great significance in public health. Therefore, this paper proposes MSMP-Net, a multi-scale neural network for end-to-end monkeypox virus skin lesion classification ConvNeXt is used as the backbone network, and designs such as inverse bottleneck layers and large convolution kernels are used to enhance the network's feature extraction capabilities. In order to effectively utilize the multi-level feature maps generated by the backbone network, a multi-scale feature fusion structure was designed. By fusing the deepest feature maps of multi-scale features, the model's ability to represent monkeypox image features is enhanced. Experimental results show that the accuracy, precision, recall, and F1-score of this method on the MSLD v2.0 dataset are 87.03 +/- 3.43%, 87.59 +/- 3.37%, 87.03 +/- 3.43%, and 86.58 +/- 3.66%, respectively.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001341380400001  
Indexed Date:  
2024-10-30  
Institution Address:  
Zhengzhou Univ Light Ind, Sch Comp Sci & Technol, Zhenzhou 450000, Peoples R China  
Institution Address:  
Zhengzhou Univ Light Ind, Sch Elect & Informat, Zhengzhou 450000, Peoples R China  
Institution:  
Institution:

### 186. Construction of a Semantic Segmentation Network for the Overhead

摘要: Accurate semantic segmentation results of the overhead catenary system (OCS) are significant for OCS component extraction and geometric parameter detection. Actually, the scenes of OCS are complex, and the density of point cloud data obtained through Light Detection and Ranging (LiDAR) scanning is uneven due to the character difference of OCS components. However, due to the inconsistent component points, it is challenging to complete better semantic segmentation of the OCS point cloud with the existing deep learning methods. Therefore, this paper proposes a point cloud multi-scale feature fusion refinement structure neural network (PMFR-Net) for semantic segmentation of the OCS point cloud. The PMFR-Net includes a prediction module and a refinement module. The innovations of the prediction module include the double efficient channel attention module (DECA) and the serial hybrid domain attention (SHDA) structure. The point cloud refinement module (PCRM) is used as the refinement module of the network. DECA focuses on detail features; SHDA strengthens the connection of contextual semantic information; PCRM further refines the segmentation results of the prediction module. In addition, this paper created and released a new dataset of the OCS point cloud. Based on this dataset, the overall accuracy (OA), F1-score, and mean intersection over union (MIoU) of PMFR-Net reached 95.77%, 93.24%, and 87.62%, respectively. Compared with four state-of-the-art (SOTA) point cloud deep learning methods, the comparative experimental results showed that PMFR-Net achieved the highest accuracy and the shortest training time. At the same time, PMFR-Net segmentation performance on S3DIS public dataset is better than the other four SOTA segmentation methods. In addition, the effectiveness of DECA, SHDA structure, and PCRM was verified in the ablation experiment. The experimental results show that this network could be applied to practical applications.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
14  
被引频次合计:  
14  
eISSN:  
2072-4292  
入藏号:  
WOS:000817365600001  
Indexed Date:  
2022-07-06  
Institution Address:  
Yangtze Univ, Sch Geosci, Wuhan 430100, Peoples R China  
Institution Address:  
Hunan Univ Sci & Technol, Hunan Prov Key Lab Geoinformat Engn Surveying Map, Xiangtan 411201, Peoples R China  
Institution Address:  
China Railway Design Corp, Tianjin 300308, Peoples R China  
Institution Address:  
Natl Petr & Nat Gas Pipeline Network Grp Co Ltd, Sci & Technol Res Inst Branch, Langfang 065000, Peoples R China  
Institution Address:  
Hunan Univ Sci & Technol, Natl Local Joint Engn Lab Geospatial Informat, Xiangtan 411201, Peoples R China  
Institution:  
Institution:  
Institution:  
China Railway Design Corp  
Institution:  
Natl Petr & Nat Gas Pipeline Network Grp Co Ltd  
Institution:

### 187. End-to-End Fovea Localisation in Colour Fundus Images With a

摘要: Accurately locating the fovea is a prerequisite for developing computer aided diagnosis (CAD) of retinal diseases. In colour fundus images of the retina, the fovea is a fuzzy region lacking prominent visual features and this makes it difficult to directly locate the fovea. While traditional methods rely on explicitly extracting image features from the surrounding structures such as the optic disc and various vessels to infer the position of the fovea, deep learning based regression technique can implicitly model the relation between the fovea and other nearby anatomical structures to determine the location of the fovea in an end-to-end fashion. Although promising, using deep learning for fovea localisation also has many unsolved challenges. In this paper, we present a new end-to-end fovea localisation method based on a hierarchical coarse-to-fine deep regression neural network. The innovative features of the new method include a multi-scale feature fusion technique and a self-attention technique to exploit location, semantic, and contextual information in an integrated framework, a multi-field-of-view (multi-FOV) feature fusion technique for context-aware feature learning and a Gaussian-shift-cropping method for augmenting effective training data. We present extensive experimental results on two public databases and show that our new method achieved state-of-the-art performances. We also present a comprehensive ablation study and analysis to demonstrate the technical soundness and effectiveness of the overall framework and its various constituent components.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
14  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
5  
被引频次合计:  
14  
ISSN:  
0278-0062  
eISSN:  
1558-254X  
入藏号:  
WOS:000604883800011  
PubMed ID:  
32915729  
Indexed Date:  
2021-01-21  
Institution Address:  
Shenzhen Univ, Coll Elect & Informat Engn, Shenzhen 518060, Peoples R China  
Institution Address:  
Shenzhen Univ, Guangdong Key Lab Intelligent Informat Proc, Shenzhen 518060, Peoples R China  
Institution Address:  
Shenzhen Inst Artificial Intelligence & Robot Soc, Shenzhen 518066, Peoples R China  
Institution Address:  
Histo Pathol Diagnost Ctr, Shanghai 201900, Peoples R China  
Institution Address:  
Wexham Pk Hosp, Frimley Hlth NHS Fdn Trust, Slough SL2 4HL, Berks, England  
Institution Address:  
Imperial Coll London, Fac Med, London SW7 2AZ, England  
Institution Address:  
Southwestern Univ Finance & Econ, Sch Econ Informat Engn, Chengdu 610074, Peoples R China  
Institution Address:  
Chengdu Everimaging Ltd, Chengdu 611800, Peoples R China  
Institution Address:  
Univ Nottingham, Sch Comp Sci, Nottingham NG8 1BB, England  
Institution:  
Institution:  
Institution:  
Institution:  
Histo Pathol Diagnost Ctr  
Institution:  
Wexham Pk Hosp  
Institution:  
Institution:  
Institution:  
Chengdu Everimaging Ltd  
Institution:

### 188. Design of Bird Sound Recognition Model Based on Lightweight

摘要: Bird sounds recognition is of great significance in bird protection. With appropriate sound classification, research can automatically predict the quality of life in the area. Nowadays, the deep learning model is used to classify bird sound data with high classification accuracy. However, the generalization ability of most existing bird sound recognition models is poor, and the complicated algorithm is applied to extract bird sound features. To address these problems, a large data set containing 264 kinds of birds is constructed in this paper to enhance the generalization ability of the model, and then a lightweight bird sound recognition model is proposed to build a lightweight feature extraction and recognition network with MobileNetV3 as the backbone. By adjusting the depthwise separable convolution in the model, the recognition ability of the model is improved. A multi-scale feature fusion structure is designed, and the Pyramid Split Attention (PSA) module is added to the multi-scale feature fusion structure to improve the adaptability of the network to scale extraction of spatial information and channel information. To improve the refinement ability of the model towards the global information, the channel attention mechanism and ordinary convolution are introduced into Bneck module which makes the Bneck module become the Bnecks module. The experimental results show that the accuracy of Top-1 and Top-5 of the model in identifying 264 kinds of birds on the self-built data set is 95.12% and 100%, which are higher than that of MobileNetV1, MobileNetV2, MobileNetV3 respectively. Although the accuracy is lower than ResNet50, the number of parameters and floating-point operations (FLOPs) of the model is only 2.6M and 127M respectively. The accuracy is only reduced by 2.25% while saving costs.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
6  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
24  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
24  
ISSN:  
2169-3536  
入藏号:  
WOS:000842054700001  
Indexed Date:  
2022-08-27  
Institution Address:  
Nanjing Forestry Univ, Coll Mech & Elect Engn, Nanjing 210037, Peoples R China  
Institution:

### 189. Attention feature fusion awareness network for vehicle target detection

摘要: Synthetic aperture radar (SAR) target detection plays a crucial role in military surveillance, earth observation, and disaster monitoring. With the development of deep learning (DL) and SAR imaging technology, numerous SAR target detection methods have been proposed and achieved better detection results. However, detecting different categories of SAR vehicle targets is still challenging due to the influence of coherent speckle noises and background clutter. This article presents a novel attention feature fusion awareness network (AFFNet) for vehicle target detection in SAR images. Specifically, we propose a multi-scale semantic attention (MSSA) module to obtain multi-scale and semantic features of target region; the variable multi-scale feature fusion (VMSFF) module is introduced to effectively fuse different feature information and alleviate target deformation interference by establishing feature correlation; the part feature awareness (PFA) module is used to obtain unique attribute of different vehicle targets to generate accurate anchor boxes. In addition, we design a candidate boundary box selection scheme, which can effectively adapt to SAR targets with different scales and categories. Overall, AFFNet is designed based on the SAR imaging mechanism and target physical feature information. To evaluate the performance of the proposed method, extensive experiments are conducted on the MSTAR dataset. The experiment results show that the proposed AFFNet obtains the mAP of 98.36 % and 97.26 % on standard operating conditions (SOCs) and extended operating conditions (EOCs), which is more efficient than the other state-of-the-art methods.  
Web of Science 核心合集中的 "被引频次":  
4  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:001054427700001  
Indexed Date:  
2023-09-04  
Institution Address:  
Xijing Univ, Sch Elect Informat, Xijing Rd, Xian 710123, Shanxi, Peoples R China  
Institution Address:  
Shandong Jianzhu Univ, Sch Surveying & Geoinformat, Jinan, Shandong, Peoples R China  
Institution Address:  
Air Force Engn Univ, Sch Informat & Nav, Xian, Shanxi, Peoples R China  
Institution:  
Institution:  
Institution:

### 190. Object detection algorithm based on image and point cloud fusion with

摘要: Object detection is the basis of autonomous driving and robot navigation. To solve the problems of insufficient information in 2D images and the large data volume, uneven density, and low detection accuracy of 3D point clouds, a new 3D object-detection network is proposed through an image and point-cloud fusion with deep learning. To reduce the calculation load, the original point cloud is first filtered with the flat interceptor corresponding to the object's frame detected in the 2D image. To address the uneven density, an improved voting model network,based on a generalized Hough transform, is proposed for multiscale feature extraction. Finally, Normal Three-Dimensional Distance Intersection over Union (N3D\_DIOU),a novel loss function, is extended from the Two-Dimensional Distance Intersection over Union (2D DIOU) loss function, which improves the consistency between the generated and target frames, and also improves the object-detection accuracy of the point cloud. Experiments on the KITTI dataset show that our algorithm improves the accuracy of three-dimensional detection by 0.71%, and the aerial-view detection accuracy by 7.28%, over outstanding classical methods.  
摘要:  
目标检测是自主驾驶和机器人导航的基础,针对二维图像信息量不足,三维点云数据量大、密度不均匀和检测精度低等问题,本文基于深度学习提出了一种融合二维图像与三维点云的目标检测网络进行三维目标检测。为减少运算量,论文首先用二维图像检测器生成的检测框对应的平截头体对原始点云进行滤波;为解决点云密度不均匀问题,提出了一种基于广义霍夫变换的改进投票模型网络用于多尺度特征提取;最后将二维DIOU(Distance Intersection over Union)损失函数扩展为三维空间的N3D\_DIOU(Normal 3 Dimensional DIOU)损失函数,提高了生成框和目标框的一致性,进一步提高了点云检测精度。在KITTI数据集上进行的大量实验表明:与经典方法相比,本文算法在汽车三维检测精度上提升了0.71%,在鸟瞰图检测精度上提升了7.28%,取得了较好效果。  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
被引频次合计:  
3  
ISSN:  
1004-924X  
入藏号:  
CSCD:7113503  
Indexed Date:  
2022-02-25  
Institution Address:  
北京交通大学机械与电子控制工程学院;;北京交通大学, ;;智慧高铁系统前沿科学中心, ;;, 北京;;北京 100044;;100044, 中国  
Institution Address:  
北京交通大学机械与电子控制工程学院, 北京 100044, 中国  
Institution Address:  
School of Mechanical,Electronic and Control Engineering,Beijing Jiaotong University;;Beijing Jiaotong University, ;;Frontiers Science Center for Smart High-speed Railway System, ;;, Beijing;;Beijing 100044;;100044  
Institution Address:  
School of Mechanical,Electronic and Control Engineering,Beijing Jiaotong University, Beijing 100044, China  
Institution:  
北京交通大学机械与电子控制工程学院;;北京交通大学  
Institution:  
北京交通大学机械与电子控制工程学院  
Institution:  
School of Mechanical,Electronic and Control Engineering,Beijing Jiaotong University;;Beijing Jiaotong University  
Institution:  
School of Mechanical,Electronic and Control Engineering,Beijing Jiaotong University

### 191. Rapid Restoration of Turbulent Degraded Images Based on Bidirectional

摘要: This study proposes a generative adversarial network (GAN) based on bidirectional multi-scale feature fusion to reconstruct target celestial images captured by various ground-based telescopes, which are influenced by atmospheric turbulence. This approach first constructs a dataset for network training by convolving a long exposure atmospheric turbulence degradation model with clear images and then validates the network's performance on a simulated turbulence image dataset. Furthermore, images of the International Space Station collected by the Munin ground-based telescope (Cassegrain-type telescope) that were influenced by atmospheric turbulence are included in this study. These images were sent to the proposed neural network model for testing. Different image restoration assessment shows that the proposed network has a good real-time performance and can produce restoration results within 0.5 s, which is more than 10 times faster than standard nonneural network restoration approaches; the peak signal to noise ratio (PSNR) is improved by 2 dB- 3 dB, and structural similarity (SSIM) is enhanced by 9.3%. Simultaneously, the proposed network has a pretty good restoration impact on degraded images that are influenced by real turbulence.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1006-4125  
入藏号:  
WOS:000892702100004  
Indexed Date:  
2022-12-23  
Institution Address:  
Chinese Acad Sci, Hefei Inst Phys Sci, Key Lab Atmospher Opt, Hefei 230031, Anhui, Peoples R China  
Institution Address:  
Univ Sci & Technol China, Hefei 230026, Anhui, Peoples R China  
Institution Address:  
Adv Laser Technol Lab Anhui Prov, Hefei 230037, Anhui, Peoples R China  
Institution:  
Institution:  
Institution:  
Adv Laser Technol Lab Anhui Prov

### 192. A new end-to-end image dehazing algorithm based on residual attention

摘要: Traditional image dehazing algorithms based on prior knowledge and deep learning rely on the atmospheric scattering model and are easy to cause color distortion and incomplete dehazing. To solve these problems,an end-to-end image dehazing algorithm based on residual attention mechanism is proposed in this paper. The network includes four modules: encoder,multi-scale feature extraction,feature fusion and decoder. The encoder module encodes the input haze image into feature map,which is convenient for subsequent feature extraction and reduces memory consumption; the multi-scale feature extraction module includes residual smoothed dilated convolution module, residual block and efficient channel attention,which can expand the receptive field and extract different scale features by filtering and weighting; the feature fusion module with efficient channel attention adjusts the channel weight dynamically,acquires rich context information and suppresses redundant information so as to enhance the ability to extract haze density image of the network; finally,the encoder module maps the fused feature nonlinearly to obtain the haze density image and then restores the haze free image. The qualitative and quantitative tests based on SOTS test set and natural haze images show good objective and subjective evaluation results. This algorithm improves the problems of color distortion and incomplete dehazing effectively.  
摘要:  
传统基于先验知识与基于学习的图像去雾算法依赖大气散射模型,容易出现颜色失真和去雾不彻底的现象。针对上述问题,提出一种端到端的基于残差注意力机制的图像去雾算法,该算法网络包括编码、多尺度特征提取、特征融合和解码4个模块。编码模块将输入的雾图编码为特征图像,便于后续特征提取并减少内存占用;多尺度特征提取模块包括残差平滑空洞卷积模块、残差块和高效通道注意力机制,能够扩大感受野并通过加权筛选提取的不同尺度特征以便融合;特征融合模块利用高效通道注意力机制,动态调整不同尺度特征的通道权重,学习丰富的上下文信息并抑制冗余信息,增强网络提取雾霾密度图像的能力从而使去雾更加彻底;解码模块对融合后的特征进行非线性映射得到雾霾密度图像,进而恢复无雾图像。通过在SOTS测试集和自然有雾图像上进行定量和定性的测试,所提方法取得了较好的客观和主观评价结果,并有效改善了颜色失真和去雾不彻底的现象。  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
1000-2758  
入藏号:  
CSCD:7146625  
Indexed Date:  
2022-05-27  
Institution Address:  
天津城建大学计算机与信息工程学院, 天津 300384, 中国  
Institution Address:  
School of Computer and Information Engineering,Tianjin Chengjian University, Tianjin 300384, China  
Institution:  
天津城建大学计算机与信息工程学院  
Institution:  
School of Computer and Information Engineering,Tianjin Chengjian University

### 193. Neural Path Planning With Multi-Scale Feature Fusion Networks

摘要: Path planning is critical for planetary rovers that perform observation and exploration missions in unknown and dangerous environment. And due to the communication delay, it is difficult for the planet rover to receive instructions from Earth in time to guide its own movement. In this work, we present a novel neural network-based algorithm to solve the global path planning problem for planetary rovers. Inspired by feature pyramid networks used for object detection, we construct a deep neural network model, termed the Pyramid Path Planning Network (P3N), which has a well-designed backbone that efficiently learns a global feature representation of the environment, and a feature pyramid branch that adaptively fuses multi-scale features from different levels to generate the local feature representation with rich semantic information. The P3N learns environmental dynamics from terrain images of planetary surface taken by satellites, without using additional elevation information to construct an explicit environmental model in advance, and can perform path planning policy after end-to-end training. We evaluate the effectiveness of the proposed method on synthetic grid maps and a realistic data set constructed from the lunar terrain images. Experimental results demonstrate that our P3N has higher prediction accuracy and faster computation speed compared to the baseline methods, and generalize better in large-scale environments.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
5  
ISSN:  
2169-3536  
入藏号:  
WOS:000886303500001  
Indexed Date:  
2022-11-29  
Institution Address:  
Dalian Maritime Univ, Sch Naval Architecture & Ocean Engn, Dalian 116026, Peoples R China  
Institution:

### 194. A Cross Working Condition Multiscale Recursive Feature Fusion Method for

摘要: As a key component of electromechanical equipment in the intelligent manufacturing process, rolling bearings play an important role to secure a safe, stable, and efficient operation. Deep learning can be used to guide a data-driven fault diagnosis which requires that all data are independently identically distribution (i.i.d). When the equipment is operated with multiple working conditions, the collected samples violates the assumption of i.i.d, which will inevitably make it difficult to extract accurate feature involved in the data. This paper proposes a deep learning based fault diagnosis model to recursively fuse the multiscale feature on cross working conditions, such that data without working condition label can also be referred to train a satisfying deep learning model for fault diagnosis of bearing operated in multiple working conditions. In the case when only a small number size of training samples for a separated working condition are available, the proposes fusion mechanism aims to establish a jointly learning mechanism between different working conditions. To verify the effectiveness of the proposed algorithm, experimental validation was performed using the Case Western Reserve University (CWRU) rolling bearing public data set. The experimental results show that the proposed method can make full use of a small amount of labeled data with working conditions and a large amount of labeled data without working conditions. In ten types of fault diagnosis tasks with different fault sizes, the fault diagnosis accuracy reaches more than 94% for 4 working conditions and more than 86% for 8 working conditions.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
8  
ISSN:  
2169-3536  
入藏号:  
WOS:000832958800001  
Indexed Date:  
2022-08-06  
Institution Address:  
Shanghai Maritime Univ, Sch Logist Engn, Shanghai 201306, Peoples R China  
Institution:

### 195. CFENet: Content-aware feature enhancement network for multi-person pose

摘要: Multi-person pose estimation is a fundamental yet challenging task in computer vision. Although great success has been made in this field due to the rapid development of deep learning, complex situations (e.g., extreme poses, occlusions, overlapped persons, and crowded scenes) are still not well solved. To further mitigate these issues, we propose a novel Content-aware Feature Enhancement Network (CFENet), which consists of three effective modules: Feature Aggregation and Selection Module (FASM), Feature Fusion Module (FFM) and Dense Upsampling Convolution (DUC) module. The FASM includes Feature Aggregation Module (FAM) and Information Selection Module (ISM). The FAM constructs the hierarchical multi-scale feature aggregations in a granular level to capture more accurate fine-grained representations. The ISM makes the aggregated representations more distinguished, which adaptively highlights the discriminative human part representations both in the spatial location and channel context. Then, we perform FFM which effectively fuses high-resolution spatial features and low-resolution semantic features to obtain more reliable context information for well-estimated joints. Finally, we adopt DUC module to generate more precise prediction, which can recover missing joint details that are usually unavailable in common upsampling process. Comprehensive experiments demonstrate that the proposed approach outperforms most of the popular methods and achieves a competitive performance with the state-of-the-art methods over three benchmark datasets: the recent big dataset CrowdPose, the COCO keypoint detection dataset and the MPII Human Pose dataset. Our code will be released upon acceptance.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000644383600002  
Indexed Date:  
2021-05-23  
Institution Address:  
Beijing Jiaotong Univ, Beijing Key Lab Traff Data Anal & Min, Beijing 100044, Peoples R China  
Institution:

### 196. HFIFNet: Hierarchical Feature Interaction Network With Multiscale Fusion

摘要: Change detection (CD) from remote sensing images has been widely used in land management and urban planning. Benefiting from deep learning, numerous methods have achieved significant results in the CD of clearly changed targets. However, there are still significant challenges in the CD of weak targets, such as targets with small size, targets with blurred boundaries, and targets with low distinguishability from the background. Feature extraction from these targets can result in the loss of critical spatial features, potentially leading to decreased CD performance. Inspired by the improvement of multiscale features for CD of weak target, a hierarchical feature interaction network with multiscale fusion was proposed. First, a hierarchical feature interactive fusion module is proposed, which achieves optimized multichannel feature interaction and enhances the distinguishability between weak targets and background. Moreover, the module also achieves cross scale feature fusion, which compensates for the loss of spatial feature of changed targets at a single scale during feature extraction. Second, VMamba Block is utilized to obtain global features, and a spatial feature localization module was proposed to enhance the saliency of spatial features such as edges and textures. The distinguishability between weak targets and irrelevant spatial features is further enhanced. Our method has been experimentally evaluated on three public datasets, and outperformed state-of-the-art approaches by 1.06%, 1.41%, and 2.63% in F1 score on the LEVIR-CD, S2Looking, and NALand datasets, respectively. These results affirm the effectiveness of our method for weak targets in CD tasks.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001411964500009  
Indexed Date:  
2025-02-18  
Institution Address:  
Univ Jinan, Sch Informat Sci & Engn, Jinan 250022, Shandong, Peoples R China  
Institution Address:  
Beihang Univ, Sch Comp Sci & Engn, Beijing 100191, Peoples R China  
Institution:  
Institution:

### 197. Automatic 3D fault segmentation based on multi-scale feature fusion

摘要: Faults serve as oil and gas storage space and transportation channels, so fault identification is significant to oil and gas exploration. Fault extraction methods based on manual identification or seismic body attributes are prone to recognition errors due to human factors or poor data quality. With the development of deep learning, researchers have proposed different network models to extract 3D faults. However, the traditional models still have room for improvement in fine-grained segmentation results and model robustness. Therefore, this study proposes a new multi-scale feature fusion network architecture named MAR-UNet. In order to solve the defect of insufficient fine granularity of traditional model segmentation results, this paper designs a local feature extraction module named Residual Sampling Convolution block (RSC block) and deploys it to MAR-UNet; at the same time, in order to improve the defect that the existing 3D model cannot effectively deal with complex spatial relationship features, this study designs a plug-and-play attention module named Mix Attention Mechanism (MAM) in the model. Finally, this paper proposes a compound loss function named Weight Focal-Dice loss for the model's weak robustness caused by sample imbalance. The results of ablation and cross-experiments show that the loss function proposed in this paper is suitable for accomplishing the fault segmentation task under the influence of sample imbalance, and the model proposed in this paper still shows good reliability and robustness when deploying the model to the actual workspace data.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1865-0473  
eISSN:  
1865-0481  
入藏号:  
WOS:001220292000002  
Indexed Date:  
2024-05-22  
Institution Address:  
Hefei Univ Technol, Sch Resources & Environm Engn, Hefei 230009, Peoples R China  
Institution:

### 198. L-Unet: A Landslide Extraction Model Using Multi-Scale Feature Fusion

摘要: At present, it is challenging to extract landslides from high-resolution remote-sensing images using deep learning. Because landslides are very complex, the accuracy of traditional extraction methods is low. To improve the efficiency and accuracy of landslide extraction, a new model is proposed based on the U-Net model to automatically extract landslides from remote-sensing images: L-Unet. The main innovations are as follows: (1) A multi-scale feature-fusion (MFF) module is added at the end of the U-Net encoding network to improve the model's ability to extract multi-scale landslide information. (2) A residual attention network is added to the U-Net model to deepen the network and improve the model's ability to represent landslide features. (3) The bilinear interpolation algorithm in the decoding network of the U-Net model is replaced by data-dependent upsampling (DUpsampling) to improve the quality of the feature maps. Experimental results showed that the precision, recall, MIoU and F1 values of the L-Unet model are 4.15%, 2.65%, 4.82% and 3.37% higher than that of the baseline U-Net model, respectively. It was proven that the new model can extract landslides accurately and effectively.  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
24  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
25  
eISSN:  
2072-4292  
入藏号:  
WOS:000808674100001  
Indexed Date:  
2022-06-20  
Institution Address:  
Hefei Univ Technol, Sch Comp & Informat, Hefei 230601, Peoples R China  
Institution Address:  
Intelligent Interconnected Syst Lab Anhui Prov, Hefei 230009, Peoples R China  
Institution Address:  
Anhui Prov Key Lab Ind Safety & Emergency Technol, Hefei 230601, Peoples R China  
Institution:  
Institution:  
Intelligent Interconnected Syst Lab Anhui Prov  
Institution:  
Anhui Prov Key Lab Ind Safety & Emergency Technol

### 199. Lightweight silkworm recognition based on Multi-scale feature fusion

摘要: Because the YOLOv4 model is unsuitable for the mobile and embedded terminals, YOLOv4 & PRIME;s lightweight MobileNetv3-YOLOv4 network significantly decreases the detection accuracy of dense silkworm targets, and the accuracy loss is too significant. A lightweight YOLOv4 detection algorithm (KM-YOLOv4) improved by multi -scale feature fusion is proposed for the target detection of dense silkworms. The Kmeans algorithm re-constructs anchor boxes suitable for different objects to enhance detection accuracy. By adding multi-scale feature fusion, the improved deep learning separable convolution MobileNetV3 lightweight backbone network replaces the YOLOv4 backbone network, reducing the computational load and model scale of the backbone network and making up for the light part of the depthwise separable convolution Accuracy loss, which improves the detection accuracy of lightweight models. The experimental results with the dense silkworm formation dataset show that the KM-YOLOv4 algorithm significantly reduces the model size by about 74% compared with the YOLOv4 algorithm and improves the detection accuracy by 1.82% with the unimproved MobileNetv3-YOLOv4 algorithm. The model can be better applied to mobile and embedded.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
11  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
16  
被引频次合计:  
18  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:000842871000002  
Indexed Date:  
2022-08-27  
Institution Address:  
Guangxi Minzu Univ, Coll Elect Informat, Nanning 530006, Peoples R China  
Institution Address:  
Guangxi Key Lab Hybrid Computat & IC Design Anal, Nanning 530006, Peoples R China  
Institution Address:  
Guangxi Univ, Agr & Anim Husb Ind Dev Res Inst, Nanning 530004, Peoples R China  
Institution Address:  
Guangxi Univ, State Key Lab Conservat & Utilizat Subtrop Agrobio, Nanning 530005, Peoples R China  
Institution Address:  
Guangxi Yizhou Ruomutianchong Sericulture Technol, Yizhou 546311, Peoples R China  
Institution:  
Institution:  
Guangxi Key Lab Hybrid Computat & IC Design Anal  
Institution:  
Institution:  
Institution:  
Guangxi Yizhou Ruomutianchong Sericulture Technol

### 200. Lightweight Landslide Detection Network for Emergency Scenarios

摘要: Landslides are geological disasters that can cause serious severe damage to properties and lead to the loss of human lives. The application of deep learning technology to optical remote sensing images can help in the detection of landslide areas. Traditional landslide detection models usually have complex structural designs to ensure accuracy. However, this complexity leads to slow detection, and these models often do not satisfy the rapid response required for the emergency monitoring of landslides. Therefore, we designed a lightweight landslide target detection network based on a CenterNet and a ResNet50 network. We replaced the BottleNeck in the backbone network of ResNet50 with a Ghost-BottleNeck structure to reduce the number of parameters in the model. We also introduced an attention mechanism module based on channel attention and spatial attention between the adjacent GhostModule modules to rich the landslide features. We introduced a lightweight multiscale fusion method in the decoding process that presented a cross-layer sampling operation for the encoding process based on Feature Pyramid Network. To down-sample from a low resolution to a high resolution and up-sample from a high resolution to a low resolution, thus skipping the medium-resolution levels in the path. We added the feature maps obtained in the previous step to the feature fusion. The Conv module that adjusts the number of channels in the multiscale feature fusion operation was replaced with the GhostModule to achieve lightweight capability. At the end of the network, we introduced a state-of-the-art Yolov5x as a teacher network for feature-based knowledge distillation to further improve the accuracy of our student network. We used challenging datasets including multiple targets and multiscale landslides in the western mountains of Sichuan, China (e.g., Danba, Jiuzhaigou, Wenchuan, and Maoxian) to evaluate the proposed lightweight landslide detection network. The experimental results show that our model satisfied landslide emergency requirements in terms of both accuracy and speed; the parameter size of the proposed lightweight model is 18.7 MB, namely, 14.6% of the size of the original CenterNet containing the ResNet50 network. The single image detection time is 52 ms-twice as fast as the original model. The detection accuracy is 76.25%, namely, 12% higher than that of the original model.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
eISSN:  
2072-4292  
入藏号:  
WOS:000940605700001  
Indexed Date:  
2023-03-18  
Institution Address:  
Minist Nat Resources, Key Lab Nat Resources Monitoring Trop & Subtrop Ar, Guangzhou 510670, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, Fac Geosci & Environm Engn, Chengdu 610032, Peoples R China  
Institution Address:  
Surveying & Mapping Inst Lands & Resource Dept Gua, Guangzhou 529001, Peoples R China  
Institution Address:  
Minist Nat Resources, Key Lab Urban Land Resources Monitoring & Simulat, Shenzhen 518034, Peoples R China  
Institution:  
Institution:  
Institution:  
Surveying & Mapping Inst Lands & Resource Dept Gua  
Institution:

### 201. Enhancing pavement health assessment: An attention-based approach for

摘要: Vehicle-based crack detection can serve as a highly effective method for assessing pavement damage. This paper proposes an automated crack detection, measurement and mapping method based on GPS tagged images acquired from a camera installed on the license plate of a vehicle. An attention-based mechanism is added to the multiscale feature fusion RetinaNet convolution neural network to detect cracks. This enables the neural network to focus on key features at different scales enhancing the model's detection capabilities. Adding the attention block effectively increases the model's performance when compared with the original network. This is followed by a lightweight categorization algorithm, a high-precision edge detection algorithm, a crack width measurement algorithm, and a mapping algorithm. The results show that the crack detection model is a well-represented model that exhibits high performance on crack-induced pavements. The process of estimating crack width is validated by achieving a low mean average relative error. The crack width is further used to assign severity to the crack. The crack detection algorithm and the crack width estimation algorithm are tested using real-world data in order to assess their effectiveness. The results of the study suggest that by integrating crack detection and width estimation, a comprehensive engineering solution for crack monitoring can be achieved. This solution can be effectively applied to conduct pavement surveys, resulting in the creation of a GPS-based map that identifies various distresses. This framework provides a holistic approach to infrastructure maintenance, allowing improved transportation infrastructure management.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
11  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001171565200001  
Indexed Date:  
2024-03-13  
Institution Address:  
Indian Inst Technol Roorkee, Dept Civil Engn, Roorkee, Uttar Pradesh, India  
Institution Address:  
Western Univ, Dept Civil & Environm Engn, London, ON, Canada  
Institution:  
Institution:

### 202. Exploring Multimodal Multiscale Features for Sentiment Analysis Using

摘要: Sentiment analysis, a challenging task in understanding human emotions expressed through diverse modalities, prompts the development of innovative solutions. Multimodal data often contains important complementary information. Effective fusion and extraction of multimodal data features are key issues in sentiment analysis. In this article, we introduce a novel sentiment analysis model that integrates multimodal multiscale features based on a fuzzy-deep neural network. First, we combine multimodal data, namely text, audio, and images, to extract intrinsic feature representations. Second, our model incorporates the fuzzy-deep neural network learning module, infused with fuzzy logic principles to enhance adaptability to the inherent vagueness in sentiment expressions. Furthermore, we integrate the dual attention mechanism that dynamically focuses on pivotal aspects within multimodal data, refining feature extraction for heightened context-awareness. Rigorous validation across three datasets, including the Multimodal Corpus of Sentiment Intensity dataset, the Multimodal Opinion Sentiment and Emotion Intensity dataset, and the Chinese Single and Multimodal Sentiment dataset, demonstrates the model's superior performance in capturing the intricacies of human emotions.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
ISSN:  
1063-6706  
eISSN:  
1941-0034  
入藏号:  
WOS:001394760500014  
Indexed Date:  
2025-01-24  
Institution Address:  
Northeastern Univ, Shenyang 110819, Peoples R China  
Institution Address:  
Peng Cheng Lab, Shenzhen 518057, Peoples R China  
Institution Address:  
Sookmyung Womens Univ, Seoul 04310, South Korea  
Institution Address:  
Nitte Meenakshi Inst Technol, Bangalore 560064, India  
Institution Address:  
SUNY Coll New Paltz, Dept Comp Sci, New Paltz, NY 12561 USA  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 203. Hyperspectral classification algorithm based on covariance pooling and

摘要: The deep convolution neural network algorithm has achieved excellent performance in hyperspectral image classification. However, these deep learning algorithms generally use first-order pooling operation, which ignores the correlation between different spectral bands. Thus, obtaining high-order statistical discriminant features is difficult. In addition, using these algorithms to choose the optimal window size and capture different receptive field information is complicated. This paper proposes a hyperspectral classification method combining covariance pooling and cross-scale feature extraction to solve the aforementioned problems. This method aims to automatically extract the complementary and discriminative information of different scales and exploit the first- and second-order pooling features to improve the classification performance. A covariance pooling and cross-scale feature extraction method is proposed for hyperspectral image classification. In this method, a cross-scale adaptive feature extraction module is designed. This module can automatically combine multiscale feature information and obtain complementary information of different visual fields, avoiding the scale selection problem. Furthermore, the first- and second-order statistics combined with spatial-spectral information are obtained using the joint pooling operation of average and fast covariance pooling. Finally, the first- and second-order pooled features are fused for classification. A total of 5%, 5%, and 1% labeled samples were randomly selected from three public hyperspectral datasets, namely, Indian pines, Houston University, and Pavia University, respectively. The overall classification accuracy of the proposed algorithm reached 97.63%, 98.48%, and 98.21%, and the classification performance was better than the state-of-the-art deep learning methods. Cross-scale feature extraction considers the complementary spatial-spectral information between different scales to obtain additional adaptive feature information. Combining fast covariance and average pooling, the discriminant features are obtained by pooling feature fusion to obtain superior classification results.  
摘要:  
深度卷积神经网络在高光谱图像分类任务上取得了优越性能。但是,主流深度学习算法通常采用一阶池化运算,容易忽略光谱之间的相关性,因而难以获取高阶统计判别特征。另外,这类算法往往难以选择最优的窗口大小去捕获不同感受野信息。针对上述问题,本文提出了一种结合协方差池化和跨尺度特征提取的高光谱影像分类方法。该方法设计了跨尺度自适应特征提取模块,能够自动提取多尺度特征,获取不同视野的互补信息,避免了尺度选择问题;进一步利用平均池化和快速协方差池化的联合池化操作,得到一阶统计量和结合空间光谱信息的二阶统计量;最终,将一阶和二阶池化特征进行融合用于分类。在3个公开高光谱数据集Indian Pines、Houston和Pavia University上分别随机选取5%、5%和1%标记样本进行训练,本文算法得到的总体分类精度分别达到97.63%、98.48%和98.21%,分类性能优于主流深度学习方法。  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
1  
ISSN:  
1007-4619  
入藏号:  
CSCD:7664727  
Indexed Date:  
2024-05-17  
Institution Address:  
湖北大学,数学与统计学学院, 应用数学湖北省重点实验室, 武汉, 湖北 430062, 中国  
Institution Address:  
湖北大学,数学与统计学学院;;宁波大学,遥感遥测产业技术研究院,宁波拾烨智能科技有限公司, 应用数学湖北省重点实验室;;, 武汉;;宁波, ;; 430062;;315211  
Institution Address:  
宁波大学,地理与空间信息技术系;;宁波大学,遥感遥测产业技术研究院,宁波拾烨智能科技有限公司, ;;, 宁波;;宁波, ;; 315211;;315211  
Institution Address:  
Faculty of Mathematics and Statistics, Hubei University, Hubei Key Laboratory of Applied Mathematics, Wuhan, Hubei 430062, China  
Institution Address:  
Faculty of Mathematics and Statistics, Hubei University;;Industrial Research Institute of Remote Sensing, Ningbo University, Shiye Intelligent Technology Co., Ltd, Hubei Key Laboratory of Applied Mathematics;;, Wuhan;;Ningbo, ;; 430062;;315211  
Institution Address:  
Department of Geography and Spatial Information Techniques, Ningbo University;;Industrial Research Institute of Remote Sensing, Ningbo University, Shiye Intelligent Technology Co., Ltd, ;;, Ningbo;;Ningbo, ;; 315211;;315211  
Institution:  
湖北大学,数学与统计学学院  
Institution:  
湖北大学,数学与统计学学院;;宁波大学,遥感遥测产业技术研究院,宁波拾烨智能科技有限公司  
Institution:  
宁波大学,地理与空间信息技术系;;宁波大学,遥感遥测产业技术研究院,宁波拾烨智能科技有限公司  
Institution:  
Faculty of Mathematics and Statistics, Hubei University  
Institution:  
Faculty of Mathematics and Statistics, Hubei University;;Industrial Research Institute of Remote Sensing, Ningbo University, Shiye Intelligent Technology Co., Ltd  
Institution:  
Department of Geography and Spatial Information Techniques, Ningbo University;;Industrial Research Institute of Remote Sensing, Ningbo University, Shiye Intelligent Technology Co., Ltd

### 204. PVNet: A Used Vehicle Pedestrian Detection Tracking and Counting Method

摘要: Advances in technology have made people's lives more prosperous. However, the increase in the number of cars and the emergence of autonomous driving technology have led to frequent road accidents. Manual observation of traffic conditions requires high labor intensity, low work efficiency, and poses safety risks. The paper proposes a deep learning-based pedestrian-vehicle detection model to replace manual observation, overcoming human resource limitations and safety concerns. The model optimizes the darknet53 backbone feature extraction network, reducing parameters and improving feature extraction capabilities, making it more suitable for pedestrian-vehicle scenarios. In addition, the PVFPN multi-scale feature fusion method is used to facilitate information exchange between different feature layers. Finally, the Bytetrack method is used for target counting and tracking. The paper model shows excellent performance in pedestrian-vehicle detection and tracking in traffic scenarios. The experimental results show that the improved model achieves a mAP@.5 of 0.952 with only 32% of the parameters compared to YOLOv8s. Furthermore, the proposed PVNet model, combined with the Bytetrack method, maintains high detection accuracy and is applicable to pedestrian-vehicle detection and tracking in traffic scenarios. In summary, this section discusses the traffic issues arising from technological development and presents the optimization and performance of the deep learning-based pedestrian-vehicle detection model, along with its potential applications in traffic scenarios.  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2071-1050  
入藏号:  
WOS:001145643900001  
Indexed Date:  
2024-02-08  
Institution Address:  
Hubei Univ Technol, Coll Comp Sci, Wuhan 430000, Peoples R China  
Institution:

### 205. Feature Refine Network for Salient Object Detection

摘要: Different feature learning strategies have enhanced performance in recent deep neural network-based salient object detection. Multi-scale strategy and residual learning strategies are two types of multi-scale learning strategies. However, there are still some problems, such as the inability to effectively utilize multi-scale feature information and the lack of fine object boundaries. We propose a feature refined network (FRNet) to overcome the problems mentioned, which includes a novel feature learning strategy that combines the multi-scale and residual learning strategies to generate the final saliency prediction. We introduce the spatial and channel 'squeeze and excitation' blocks (scSE) at the side outputs of the backbone. It allows the network to concentrate more on saliency regions at various scales. Then, we propose the adaptive feature fusion module (AFFM), which efficiently fuses multi-scale feature information in order to predict superior saliency maps. Finally, to supervise network learning of more information on object boundaries, we propose a hybrid loss that contains four fundamental losses and combines properties of diverse losses. Comprehensive experiments demonstrate the effectiveness of the FRNet on five datasets, with competitive results when compared to other relevant approaches.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
1424-8220  
入藏号:  
WOS:000818251100001  
PubMed ID:  
35746271  
Indexed Date:  
2022-07-06  
Institution Address:  
Xinjiang Univ, Coll Informat Sci & Engn, Urumqi 830046, Peoples R China  
Institution:

### 206. Deep 3D multi-scale dual path network for automatic lung nodule

摘要: Lung cancer is the cancer with the highest mortality rate in the USA. Computed tomography (CT) scans for early diagnosis of pulmonary nodules can detect lung cancer in time. To overcome the limitations of the segmentation and handcrafted features required by traditional methods, we take deep neural network to diagnose lung cancer. In this work, we propose a deep end-to-end 3D multi-scale network based on dual path architecture (3D MS-DPN) for lung nodule classification. The 3D MS-DPN model incorporates the dual path architecture to reduce the complexity and improve the accuracy of the model fully considering the 3D nature of CT scan while performing 3D convolution. Meanwhile, the multi-scale feature fusion is used to eliminate the effects which the size of lung nodules varied widely and nodules occupying few regions and slices in CT scan. Our model achieves competitive performance on the LIDC-IDRI dataset compared to the recent related works.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1752-6418  
eISSN:  
1752-6426  
入藏号:  
WOS:000823172500003  
Indexed Date:  
2023-04-02  
Institution Address:  
Jilin Univ, Coll Comp Sci & Technol, Changchun, Peoples R China  
Institution Address:  
Jilin Univ, Coll Software, Changchun, Peoples R China  
Institution:  
Institution:

### 207. Multiscale Hybrid Convolutional Deep Neural Networks with Channel

摘要: Attention mechanisms can improve the performance of neural networks, but the recent attention networks bring a greater computational overhead while improving network performance. How to maintain model performance while reducing complexity is a hot research topic. In this paper, a lightweight Mixture Attention (MA) module is proposed to improve network performance and reduce the complexity of the model. Firstly, the MA module uses multi-branch architecture to process the input feature map in order to extract the multi-scale feature information of the input image. Secondly, in order to reduce the number of parameters, each branch uses group convolution independently, and the feature maps extracted by different branches are fused along the channel dimension. Finally, the fused feature maps are processed using the channel attention module to extract statistical information on the channels. The proposed method is efficient yet effective, e.g., the network parameters and computational cost are reduced by 9.86% and 7.83%, respectively, and the Top-1 performance is improved by 1.99% compared with ResNet50. Experimental results on common-used benchmarks, including CIFAR-10 for classification and PASCAL-VOC for object detection, demonstrate that the proposed MA outperforms the current SOTA methods significantly by achieving higher accuracy while having lower model complexity.  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
1099-4300  
入藏号:  
WOS:000856419500001  
PubMed ID:  
36141066  
Indexed Date:  
2022-09-27  
Institution Address:  
Guizhou Univ, Elect Engn Coll, Guiyang 550025, Peoples R China  
Institution Address:  
Power China Guizhou Engn Co Ltd, Guiyang 550001, Peoples R China  
Institution:  
Institution:  
Power China Guizhou Engn Co Ltd

### 208. MCF-YOLOv5: A Small Target Detection Algorithm Based on Multi-Scale

摘要: In recent years, many deep learning-based object detection methods have performed well in various applications, especially in large-scale object detection. However, when detecting small targets, previous object detection algorithms cannot achieve good results due to the characteristics of the small targets themselves. To address the aforementioned issues, we propose the small object algorithm model MCF-YOLOv5, which has undergone three improvements based on YOLOv5. Firstly, a data augmentation strategy combining Mixup and Mosaic is used to increase the number of small targets in the image and reduce the interference of noise and changes in detection. Secondly, in order to accurately locate the position of small targets and reduce the impact of unimportant information on small targets in the image, the attention mechanism coordinate attention is introduced in YOLOv5's neck network. Finally, we improve the Feature Pyramid Network (FPN) structure and add a small object detection layer to enhance the feature extraction ability of small objects and improve the detection accuracy of small objects. The experimental results show that, with a small increase in computational complexity, the proposed MCF-YOLOv5 achieves better performance than the baseline on both the VisDrone2021 dataset and the Tsinghua Tencent100K dataset. Compared with YOLOv5, MCF-YOLOv5 has improved detection APsmall by 3.3% and 3.6%, respectively.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2078-2489  
入藏号:  
WOS:001232741400001  
Indexed Date:  
2024-06-03  
Institution Address:  
Shandong Univ Technol, Sch Mech Engn, Zibo 255000, Peoples R China  
Institution:

### 209. RCFNC: a resolution and contrast fusion network with ConvLSTM for

摘要: Low-light image enhancement based on deep learning has achieved breakthroughs recently. However, the current methods based on deep learning have problems with inadequate resolution enhancement or inadequate contrast. To address these problems, this paper proposes a resolution and contrast fusion network with ConvLSTM (RCFNC) for low-light image enhancement. The network is mainly constructed by four parts, including resolution enhancement branch, contrast enhancement branch, multi-scale feature fusion block (MFFB), and convolution long short-time memory block (ConvLSTM). Specifically, to improve the resolution of the low-light image, a resolution enhancement branch consisting of multi-scale differential feature blocks is proposed, using residual features at different scales to enhance the spatial details of image. To enhance the contrast of the image, a contrast enhancement branch consisting of adaptive convolution residual blocks is introduced to learn the mapping relationship between global and local features in the image. In addition, a weighted fusion is performed using MFFB to better balance the resolution and contrast features obtained from the above branches. Finally, to improve the learning capability of the model, ConvLSTM is added to filter redundant information. Experiments on the LOL, MIT5K, and five benchmark datasets show that RCFNC outperforms current state-of-the-art methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001024240700002  
Indexed Date:  
2023-07-24  
Institution Address:  
Zhengzhou Univ Light Ind, Sch Comp & Commun Engn, Zhengzhou, Peoples R China  
Institution:

### 210. DA-IMRN: Dual-Attention-Guided Interactive Multi-Scale Residual Network

摘要: Deep learning-based fusion of spectral-spatial information is increasingly dominant for hyperspectral image (HSI) classification. However, due to insufficient samples, current feature fusion methods often neglect joint interactions. In this paper, to further improve the classification accuracy, we propose a dual-attention-guided interactive multi-scale residual network (DA-IMRN) to explore the joint spectral-spatial information and assign pixel-wise labels for HSIs without information leakage. In DA-IMRN, two branches focusing on spatial and spectral information separately are employed for feature extraction. A bidirectional-attention mechanism is employed to guide the interactive feature learning between two branches and promote refined feature maps. In addition, we extract deep multi-scale features corresponding to multiple receptive fields from limited samples via a multi-scale spectral/spatial residual block, to improve classification performance. Experimental results on three benchmark datasets (i.e., Salinas Valley, Pavia University, and Indian Pines) support that attention-guided multi-scale feature learning can effectively explore the joint spectral-spatial information. The proposed method outperforms state-of-the-art methods with the overall accuracy of 91.26%, 93.33%, and 82.38%, and the average accuracy of 94.22%, 89.61%, and 80.35%, respectively.  
Web of Science 核心合集中的 "被引频次":  
8  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
9  
eISSN:  
2072-4292  
入藏号:  
WOS:000755647700001  
Indexed Date:  
2022-02-25  
Institution Address:  
China Univ Min & Technol, Minist Educ, Engn Res Ctr Intelligent Control Underground Spac, Xuzhou 221116, Jiangsu, Peoples R China  
Institution Address:  
China Univ Min & Technol, Sch Environm Sci & Spatial Informat, Xuzhou 221116, Jiangsu, Peoples R China  
Institution Address:  
Univ Derby, Sch Elect Comp & Math, Derby DE22 1GB, England  
Institution Address:  
Univ British Columbia, Dept Elect & Comp Engn, Vancouver, BC V6T 1Z4, Canada  
Institution:  
Institution:  
Institution:  
Institution:

### 211. Intelligent identification and segmentation method of wellbore fractures

摘要: In view of the huge workload, strong subjectivity in artificial identification, and poor stability in the fracture identification and processing of logging images, this paper introduces the computer vision technology and deep learning framework into the analysis and interpretation of logging images, builds an intelligent identification and segmentation network model of fracture morphology, and intelligently identifies wellbore fractures in resistivity logging images. First, the model extracts the shallow and deep features of wellbore images through multi-scale dilated convolution and attention mechanism, and multi-scale fusion of shallow and deep features is conducted to form new features with more representation ability. According to the new features, the two pixel classification is carried out to complete foreground and background type identification of each pixel in the logging images. Several pixels classified as foreground present the contour of the fractured area. The multi-scale feature fusion model can fully retain more contour details of the fracture image from the micro perspective, and the identification and classification accuracy of each fracture pixel reaches almost 80.0%. Finally, by drawing lessons from the evaluation system of human eye visual similarity, a performance evaluation algorithm is designed for intelligently identifying fracture contour from the macro perspective. The evaluation results show that when the visual similarity perception rating is grade II, 81.3% and 80.0% of identification results in the fracture region in the training set and test set images are basically consistent with the artificial identification results. The results indicate that the proposed method can replace artificial interpretation to complete fracture identification and marking, greatly reduce the image analysis workload and carefully outline the fracture contour. Meanwhile, it is conducive to the rapid and timely judgment of wellbore and shaft stability, thus providing technical support for subsequent intelligent quantitative evaluation and calculation of fractured areas.  
摘要:  
目前测井图像裂缝识别、处理过程工作量巨大,人工识别主观性强、稳定性差,为此,提出将计算机视觉技术和深度学习框架引入测井成像图分析解读领域,构建新型裂缝形态智能识别网络模型,实现了电阻率测井成像图中井壁裂缝区域的智能识别与分割标注。首先,通过多尺度空洞卷积结合注意力机制提取电阻率测井井壁成像图中浅层和深层特征,并将深、浅层特征进行多尺度融合,形成更具表征能力的新特征。然后,根据该特征进行像素点二分类,完成每个像素点的前景、背景类型识别,若干个前景分类的像素点对应裂缝区域的轮廓。多尺度特征融合模型从微观角度充分保留了裂缝区域图像轮廓细节,裂缝区域关联像素点识别分类准确率接近80%。最后,进一步借鉴人眼视觉相似度评价体系,从宏观角度设计裂缝轮廓智能识别性能评价算法。评价结果表明,当视觉相似度感受评级为Ⅱ级时,训练集和测试集图像中与人工识别结果基本一致的裂缝区域分别达到81.3%和80.0%,说明所提方法可替代人工解释完成裂缝的识别和标注工作,能大幅减少图像分析工作量,细致勾勒出裂缝区域轮廓线。同时,有利于及时、迅速地判断井筒、井壁稳定性,为后续裂缝区域的智能定量评价、计算提供技术支撑。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1000-7210  
入藏号:  
CSCD:7604007  
Indexed Date:  
2024-02-02  
Institution Address:  
西南石油大学电气信息学院, 成都, 四川 610500, 中国  
Institution Address:  
中国石化石油工程技术研究院, 北京 102200, 中国  
Institution Address:  
School of Electrical Engineering and Information, Southwest Petroleum University, Chengdu, Sichuan 610500, China  
Institution Address:  
SINOPEC Petroleum Engineering Technology Research Institute, Beijing 102200, China  
Institution:  
西南石油大学电气信息学院  
Institution:  
中国石化石油工程技术研究院  
Institution:  
School of Electrical Engineering and Information, Southwest Petroleum University  
Institution:  
SINOPEC Petroleum Engineering Technology Research Institute

### 212. Neural style transfer combined with EfficientDet for thermal

摘要: Hindrance caused while performing object detection during border or perimeter surveillance at night due to low resolution and atmospheric noise inside infrared frames, demands for a sophisticated detection framework. To deal with the challenges, we present a novel deep learning framework dedicated completely for automated thermal surveillance. The framework is a composition of two major modules. A novel style transferred enhanced image module STEIM enhances the resolution of IR input frame by utilizing the contextual information and exploiting the local and global features to preserve the high frequency details. The enhanced IR frame is further fed inside the fine-tuned EfficientDet module EDM that comprises weighted two-way feature network. It has an advantage of effectual multi-scale feature fusion contributing to accurate and efficient object detection. We have obtained mAP of 92.83% with 94.07% accuracy running at 95fps on FLIR test frames and mAP of 87.51% with 88.53% accuracy running at 89fps on OTCVBS test frames. Experimental results on two benchmark datasets FLIR and OTCBVS completely demonstrate suitability of the proposed framework for automatic surveillance using thermal imaging systems.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
10  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:000686026700002  
Indexed Date:  
2021-08-27  
Institution Address:  
Punjab Engn Coll, Dept Elect & Commun Engn, Sect 12, Chandigarh, India  
Institution:

### 213. Multi-Scale Feature Fusion for Interior Style Detection

摘要: Text-based search engines can extract various types of information when a user enters an appropriate search query. However, a text-based search often fails in image retrieval when image understanding is needed. Deep learning (DL) is often used for image task problems, and various DL methods have successfully extracted visual features. However, as human perception differs for each individual, a dataset with an abundant number of images evaluated by human subjects is not available in many cases, although DL requires a considerable amount of data to estimate space ambiance, and the DL models that have been created are difficult to understand. In addition, it has been reported that texture is deeply related to space ambiance. Therefore, in this study, bag of visual words (BoVW) is used. By applying a hierarchical representation to BoVW, we propose a new interior style detection method using multi-scale features and boosting. The multi-scale features are created by combining global features from BoVW and local features that use object detection. Experiments on an image understanding task were conducted on a dataset consisting of room images with multiple styles. The results show that the proposed method improves the accuracy by 0.128 compared with the conventional method and by 0.021 compared with a residual network. Therefore, the proposed method can better detect interior style using multi-scale features.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2076-3417  
入藏号:  
WOS:000866625100001  
Indexed Date:  
2022-10-21  
Institution Address:  
Doshisha Univ, Grad Sch Sci & Engn, 1-3 Tatara Miyakodani, Kyotanabe, Kyoto 6100394, Japan  
Institution Address:  
Doshisha Univ, Dept Sci & Engn, 1-3 Tatara Miyakodani, Kyotanabe, Kyoto 6100394, Japan  
Institution:  
Institution:

### 214. DPSSD: Dual-Path Single-Shot Detector

摘要: Object detection is one of the most important and challenging branches of computer vision. It has been widely used in people's lives, such as for surveillance security and autonomous driving. We propose a novel dual-path multi-scale object detection paradigm in order to extract more abundant feature information for the object detection task and optimize the multi-scale object detection problem, and based on this, we design a single-stage general object detection algorithm called Dual-Path Single-Shot Detector (DPSSD). The dual path ensures that shallow features, i.e., residual path and concatenation path, can be more easily utilized to improve detection accuracy. Our improved dual-path network is more adaptable to multi-scale object detection tasks, and we combine it with the feature fusion module to generate a multi-scale feature learning paradigm called the "Dual-Path Feature Pyramid". We trained the models on PASCAL VOC datasets and COCO datasets with 320 pixels and 512 pixels input, respectively, and performed inference experiments to validate the structures in the neural network. The experimental results show that our algorithm has an advantage over anchor-based single-stage object detection algorithms and achieves an advanced level in average accuracy. Researchers can replicate the reported results of this paper.  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:000817575900001  
PubMed ID:  
35746398  
Indexed Date:  
2022-07-06  
Institution Address:  
Qilu Univ Technol, Sch Mech Engn, Shandong Acad Sci, Jinan 250300, Peoples R China  
Institution Address:  
Qilu Univ Technol, Sch Informat & Automat Engn, Shandong Acad Sci, Jinan 250300, Peoples R China  
Institution:  
Institution:

### 215. YOLOv8s-Longan: a lightweight detection method for the longan

摘要: Introduction Due to the limited computing power and fast flight speed of the picking of unmanned aerial vehicles (UAVs), it is important to design a quick and accurate detecting algorithm to obtain the fruit position.Methods This paper proposes a lightweight deep learning algorithm, named YOLOv8s-Longan, to improve the detection accuracy and reduce the number of model parameters for fruitpicking UAVs. To make the network lightweight and improve its generalization performance, the Average and Max pooling attention (AMA) attention module is designed and integrated into the DenseAMA and C2f-Faster-AMA modules on the proposed backbone network. To improve the detection accuracy, a crossstage local network structure VOVGSCSPC module is designed, which can help the model better understand the information of the image through multiscale feature fusion and improve the perception and expression ability of the model. Meanwhile, the novel Inner-SIoU loss function is proposed as the loss function of the target bounding box.Results and discussion The experimental results show that the proposed algorithm has good detection ability for densely distributed and mutually occluded longan string fruit under complex backgrounds with a mAP@0.5 of 84.3%. Compared with other YOLOv8 models, the improved model of mAP@0.5 improves by 3.9% and reduces the number of parameters by 20.3%. It satisfies the high accuracy and fast detection requirements for fruit detection in fruit-picking UAV scenarios.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1664-462X  
入藏号:  
WOS:001413106000001  
PubMed ID:  
39911656  
Indexed Date:  
2025-02-09  
Institution Address:  
South China Agr Univ, Coll Engn, Guangzhou, Peoples R China  
Institution Address:  
Guangdong Lab Lingnan Modern Agr, Guangzhou, Peoples R China  
Institution Address:  
State Key Lab Agr Equipment Technol, Beijing, Peoples R China  
Institution:  
Institution:  
Institution:  
State Key Lab Agr Equipment Technol

### 216. Hyperspectral Image Denoising via Adversarial Learning

摘要: Due to sensor instability and atmospheric interference, hyperspectral images (HSIs) often suffer from different kinds of noise which degrade the performance of downstream tasks. Therefore, HSI denoising has become an essential part of HSI preprocessing. Traditional methods tend to tackle one specific type of noise and remove it iteratively, resulting in drawbacks including inefficiency when dealing with mixed noise. Most recently, deep neural network-based models, especially generative adversarial networks, have demonstrated promising performance in generic image denoising. However, in contrast to generic RGB images, HSIs often possess abundant spectral information; thus, it is non-trivial to design a denoising network to effectively explore both spatial and spectral characteristics simultaneously. To address the above issues, in this paper, we propose an end-to-end HSI denoising model via adversarial learning. More specifically, to capture the subtle noise distribution from both spatial and spectral dimensions, we designed a Residual Spatial-Spectral Module (RSSM) and embed it in an UNet-like structure as the generator to obtain clean images. To distinguish the real image from the generated one, we designed a discriminator based on the Multiscale Feature Fusion Module (MFFM) to further improve the quality of the denoising results. The generator was trained with joint loss functions, including reconstruction loss, structural loss and adversarial loss. Moreover, considering the lack of publicly available training data for the HSI denoising task, we collected an additional benchmark dataset denoted as the Shandong Feicheng Denoising (SFD) dataset. We evaluated five types of mixed noise across several datasets in comparative experiments, and comprehensive experimental results on both simulated and real data demonstrate that the proposed model achieves competitive results against state-of-the-art methods. For ablation studies, we investigated the structure of the generator as well as the training process with joint losses and different amounts of training data, further validating the rationality and effectiveness of the proposed method.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
25  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
25  
eISSN:  
2072-4292  
入藏号:  
WOS:000787407600001  
Indexed Date:  
2022-05-03  
Institution Address:  
Shanghai Univ, Shanghai Inst Adv Commun & Data Sci, Joint Int Res Lab Specialty Fiber Opt & Adv Commu, Key Lab Specialty Fiber Opt & Opt Access Networ, Shanghai 200444, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Hangzhou Inst Adv Study, Hangzhou 310024, Peoples R China  
Institution:  
Institution:

### 217. Young-gaze: an appearance-based gaze estimation solution for adolescents

摘要: According to the World Health Organization survey, the global incidence of adolescent mental illness is 28%, while the disease detection rate is only 24.6%. Many existing works use complex eye-tracking devices to study adolescent autism, depression, and other mental illness. In this paper, we propose a gaze estimation method to replace eye-tracking devices. Appearance-based methods with deep learning can predict the point of gaze by using a monocular camera, which requires a large number of samples to learn. However, the samples collected in publicly available gaze estimation datasets are mainly adults and not adolescents. To address the above issue, our work makes two contributions. First, we collected images from 107 adolescents aged 10-14 years by laptops under uncontrolled conditions to create the Young-Gaze dataset. Second, we propose a Multi-scale Feature Fusion-based Calibration Network (MFFC-Net) to deeply fuse the eye-face features for gaze estimation. The proposed MFFC-Net achieves the better performance on Young-Gaze and other public datasets.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001262883500003  
Indexed Date:  
2024-07-14  
Institution Address:  
Shanghai Univ, Sch Commun & Informat Engn, Shanghai 200444, Peoples R China  
Institution Address:  
Shanghai Univ, Wenzhou Inst, Wenzhou 325000, Peoples R China  
Institution Address:  
Montverde Future Acad, Shanghai 201103, Peoples R China  
Institution:  
Institution:  
Institution:  
Montverde Future Acad

### 218. A context-aware progressive attention aggregation network for fabric

摘要: Fabric defect detection plays a critical role for measuring quality control in the textile manufacturing industry. Deep learning-based saliency models can quickly spot the most interesting regions that attract human attention from the complex background, which have been successfully applied in fabric defect detection. However, most of the previous methods mainly adopted multi-level feature aggregation yet ignored the complementary relationship among different features, and thus resulted in poor representation capability for the tiny and slender defects. To remedy these issues, we propose a novel saliency-based fabric defect detection network, which can exploit the complementary information between different layers to enhance the representation features ability and discrimination of defects. Specifically, a multi-scale feature aggregation unit (MFAU) is proposed to effectively characterize the multi-scale contextual features. Besides, a feature fusion refinement module (FFR) composed of an attention fusion unit (AFU) and an auxiliary refinement unit (ARU) is designed to exploit complementary important information and further refine the input features for enhancing the discriminative ability of defect features. Finally, a multi-level deep supervision (MDS) is adopted to guide the model to generate more accurate saliency maps. Under different evaluation metrics, our proposed method outperforms most state-of-the-art methods on our developed fabric datasets.  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1558-9250  
入藏号:  
WOS:001007680900001  
Indexed Date:  
2023-06-24  
Institution Address:  
Zhongyuan Univ Technol, Sch Elect & Informat Engn, Zhengzhou, Peoples R China  
Institution Address:  
South Campus Zhongyuan Inst Technol, 1 Huaihe Rd, Zhengzhou 450007, Henan Province, Peoples R China  
Institution:  
Institution:  
South Campus Zhongyuan Inst Technol

### 219. A Lightweight SAR Ship Detection Network Based on Deep Multiscale

摘要: Deep learning has proven to be highly effective in synthetic aperture radar (SAR) image target detection. However, many latest deep learning models have predominantly focused on increasing depth and size to enhance detection accuracy, often ignoring the balance between accuracy and detection speed, as well as the practical deployment of these models on hardware platforms. Therefore, a lightweight algorithm for SAR ship detection is designed in this article. First, we propose a preliminary lightweight scheme, including a multiscale feature learning augmented backbone, a lightweight feature fusion neck, and a parameter-sharing lightweight detection head. Second, unimportant branches of the network are pruned to further compress the model. Finally, the detection accuracy of the model is enhanced by knowledge distillation without augmenting the model volume, which compensates for the accuracy loss caused by model compression. Experimental validation is conducted on three SAR image ship detection datasets (SSDD, high-resolution SAR images dataset, large-scale SAR ship detection dataset-v1.0) to thoroughly assess the effectiveness of the proposed lightweight algorithm. Experimental results on the three datasets demonstrate that the proposed method achieves a model volume reduction to one-third of the baseline while maintaining a minimal decrease in detection accuracy. In SSDD, the proposed method achieved 98.7 accuracy, 0.92M parameters, 3.1G FLOPS and 2.1 MB size of 1.5X pruning rate. Furthermore, it outperforms other state-of-the-art lightweight detectors.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001385590900009  
Indexed Date:  
2025-01-03  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Informat & Commun Engn, State Key Lab Networking & Switching Technol, Beijing 100876, Peoples R China  
Institution:

### 220. Multilevel Context Feature Fusion for Semantic Segmentation of ALS Point

摘要: Semantic segmentation of airborne laser scanning (ALS) point clouds using deep learning is a hot research in remote sensing and photogrammetry. A current trend is to aggregate contextual features from different scales for boosting network generalization and diversity discrimination capabilities. One main challenge is how to achieve effective fusion with multi-scale information. In this letter, we propose a multilevel context feature fusion network (MCFN) for semantic segmentation of ALS point cloud based on an encoder-decoder structure. More specifically, we design the squeeze-expansion shared multilayer perceptron (SE-MLP) module following kernel point convolution (KPConv) in the encoding stage, which can extend the receptive field of KPConv. To aggregate low-level features and highlevel representations, we establish channel self-attention between skip connections. In the decoding stage, we develop a crosslayer attention fusion (CAF) module to generate additional discriminative channel features by fusing multiscale features at different upsampling layers. Experiments on the ISPRS and LASDU datasets demonstrate the superiority of the proposed method. Code: https://github.com/SC-shendazt/MCFN.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
12  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
12  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001037848500010  
Indexed Date:  
2023-08-20  
Institution Address:  
Xian Technol Univ, Sch Mechatron Engn, Xian 710054, Peoples R China  
Institution Address:  
Chongqing Univ, Coll Comp Sci, Chongqing 400044, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Sch Commun & Informat Engn, Chongqing 400065, Peoples R China  
Institution Address:  
Univ Sydney, Fac Engn, Camperdown, NSW 2002, Australia  
Institution:  
Institution:  
Institution:  
Institution:

### 221. Prediction of drug-target binding affinity based on multi-scale feature

摘要: Accurate prediction of drug-target binding affinity (DTA) plays a pivotal role in drug discovery and repositioning. Although deep learning methods are widely used in DTA prediction, two significant challenges persist: (i) how to effectively represent the complex structural information of proteins and drugs; (ii) how to precisely model the mutual interactions between protein binding sites and key drug substructures. To address these challenges, we propose a MSFFDTA (Multi-scale feature fusion for predicting drug target affinity) model, in which multi-scale encoders effectively capture multi-level structural information of drugs and proteins are designed. And then a Selective Cross Attention (SCA) mechanism is developed to filter out the trivial interactions between drug-protein substructure pairs and retain the important ones, which will make the proposed model better focusing on these key interactions and offering insights into their underlying mechanism. Experimental results on two benchmark datasets demonstrate that MSFFDTA is superior to several state-of-the-art methods across almost all comparison metrics. Finally, we provide the ablation and case studies with visualizations to verify the effectiveness and the interpretability of MSFFDTA. The source code is freely available at https://github.com/whitehat32/MSFF-DTA/.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
1  
eISSN:  
1879-0534  
入藏号:  
MEDLINE:38870725  
PubMed ID:  
38870725  
Indexed Date:  
2024-06-18  
Institution Address:  
School of Computer Science, Northwestern Polytechnical University, Xi'an, 710072, China. Electronic address: huiyu@nwpu.edu.cn.  
Institution Address:  
School of Computer Science, Northwestern Polytechnical University, Xi'an, 710072, China. Electronic address: xwx@mail.nwpu.edu.cn.  
Institution Address:  
School of Computer Science, Northwestern Polytechnical University, Xi'an, 710072, China. Electronic address: 1813163384@mail.nwpu.edu.cn.  
Institution Address:  
School of Computer Science, Northwestern Polytechnical University, Xi'an, 710072, China. Electronic address: liuzun@nwpu.edu.cn.  
Institution Address:  
School of Life Sciences, Northwestern Polytechnical University, Xi'an, 710072, China. Electronic address: jianyushi@nwpu.edu.cn.  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 222. Rapid Target Detection of Fruit Trees Using UAV Imaging and Improved

摘要: The detection and counting of fruit tree canopies are important for orchard management, yield estimation, and phenotypic analysis. Previous research has shown that most fruit tree canopy detection methods are based on the use of traditional computer vision algorithms or machine learning methods to extract shallow features such as color and contour, with good results. However, due to the lack of robustness of these features, most methods are hardly adequate for the recognition and counting of fruit tree canopies in natural scenes. Other studies have shown that deep learning methods can be used to perform canopy detection. However, the adhesion and occlusion of fruit tree canopies, as well as background noise, limit the accuracy of detection. Therefore, to improve the accuracy of fruit tree canopy recognition and counting in real-world scenarios, an improved YOLOv4 (you only look once v4) is proposed, using a dataset produced from fruit tree canopy UAV imagery, combined with the Mobilenetv3 network, which can lighten the model and increase the detection speed, combined with the CBAM (convolutional block attention module), which can increase the feature extraction capability of the network, and combined with ASFF (adaptively spatial feature fusion), which enhances the multi-scale feature fusion capability of the network. In addition, the K-means algorithm and linear scale scaling are used to optimize the generation of pre-selected boxes, and the learning strategy of cosine annealing is combined to train the model, thus accelerating the training speed of the model and improving the detection accuracy. The results show that the improved YOLOv4 model can effectively overcome the noise in an orchard environment and achieve fast and accurate recognition and counting of fruit tree crowns while lightweight the model. The mAP reached 98.21%, FPS reached 96.25 and F1-score reached 93.60% for canopy detection, with a significant reduction in model size; the average overall accuracy (AOA) reached 96.73% for counting. In conclusion, the YOLOv4-Mobilenetv3-CBAM-ASFF-P model meets the practical requirements of orchard fruit tree canopy detection and counting in this study, providing optional technical support for the digitalization, refinement, and smart development of smart orchards.  
在中国科学引文数据库中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
4  
Web of Science 核心合集中的 "被引频次":  
20  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
23  
eISSN:  
2072-4292  
入藏号:  
WOS:000851856800001  
Indexed Date:  
2022-09-14  
Institution Address:  
Zhejiang A&F Univ, Coll Math & Comp Sci, Hangzhou 311300, Peoples R China  
Institution Address:  
Zhejiang Univ, Coll BioSyst Engn & Food Sci, Hangzhou 310058, Peoples R China  
Institution:  
Institution:

### 223. Optical-to-SAR Translation Based on CDA-GAN for High-Quality Training

摘要: Abundant datasets are critical to train models based on deep learning technologies for ship detection applications. Compared with optical images, ship detection based on synthetic aperture radar (SAR) (especially the high-Earth-orbit spaceborne SAR launched recently) lacks enough training samples. A novel cross-domain attention GAN (CDA-GAN) model is proposed for optical-to-SAR translation, which can generate high-quality SAR amplitude training samples of a target by optical image conversion. This high quality includes high geometry structure similarity of the target compared with the corresponding optical image and low background noise around the target. In the proposed model, the cross-domain attention mechanism and cross-domain multi-scale feature fusion are designed to improve the quality of samples for detection based on the generative adversarial network (GAN). Specifically, a cross-domain attention mechanism is designed to simultaneously emphasize discriminative features from optical images and SAR images at the same time. Moreover, a designed cross-domain multi-scale feature fusion module further emphasizes the geometric information and semantic information of the target in a feature graph from the perspective of global features. Finally, a reference loss is introduced in CDA-GAN to completely retain the extra features generated by the cross-domain attention mechanism and cross-domain multi-scale feature fusion module. Experimental results demonstrate that the training samples generated by the proposed CDA-GAN can obtain higher ship detection accuracy using real SAR data than the other state-of-the-art methods. The proposed method is generally available for different orbit SARs and can be extended to the high-Earth-orbit spaceborne SAR case.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:001305916400001  
Indexed Date:  
2024-09-15  
Institution Address:  
Northwestern Polytech Univ, Sch Elect & Informat, Xian 710072, Peoples R China  
Institution Address:  
Cent South Univ, Sch Automat, Changsha 410083, Peoples R China  
Institution:  
Institution:

### 224. Semi-Supervised Remote Sensing Building Change Detection with Joint

摘要: The timely updating of the spatial distribution of buildings is essential to understanding a city's development. Deep learning methods have remarkable benefits in quickly and accurately recognizing these changes. Current semi-supervised change detection (SSCD) methods have effectively reduced the reliance on labeled data. However, these methods primarily focus on utilizing unlabeled data through various training strategies, neglecting the impact of pseudo-changes and learning bias in models. When dealing with limited labeled data, abundant low-quality pseudo-labels generated by poorly performing models can hinder effective performance improvement, leading to the incomplete recognition results of changes to buildings. To address this issue, we propose a feature multi-scale information interaction and complementation semi-supervised method based on consistency regularization (MSFG-SemiCD), which includes a multi-scale feature fusion-guided change detection network (MSFGNet) and a semi-supervised update method. Among them, the network facilitates the generation of multi-scale change features, integrates features, and captures multi-scale change targets through the temporal difference guidance module, the full-scale feature fusion module, and the depth feature guidance fusion module. Moreover, this enables the fusion and complementation of information between features, resulting in more complete change features. The semi-supervised update method employs a weak-to-strong consistency framework to achieve model parameter updates while maintaining perturbation invariance of unlabeled data at both input and encoder output features. Experimental results on the WHU-CD and LEVIR-CD datasets confirm the efficacy of the proposed method. There is a notable improvement in performance at both the 1% and 5% levels. The IOU in the WHU-CD dataset increased by 5.72% and 6.84%, respectively, while in the LEVIR-CD dataset, it improved by 18.44% and 5.52%, respectively.  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:001323400400001  
Indexed Date:  
2024-10-07  
Institution Address:  
China Univ Geosci, Sch Comp Sci, Wuhan 430074, Peoples R China  
Institution Address:  
Minist Educ, Engn Res Ctr Nat Resource Informat Management & Di, Wuhan 430074, Peoples R China  
Institution Address:  
China Univ Geosci Wuhan, Key Lab Geol Survey & Evaluat, Minist Educ, Wuhan 430079, Peoples R China  
Institution Address:  
China Univ Geosci, Sch Geog & Informat Engn, Wuhan 430074, Peoples R China  
Institution:  
Institution:  
Minist Educ  
Institution:  
Institution:

### 225. Camouflage Object Detection Based on Feature Fusion and Edge Detection

摘要: Camouflaged Object Detection (COD) holds significant research and application value in various fields. The ability of deep learning is pushing the performance of target detection algorithms to new heights. Designing a network that effectively integrates features of different layer sizes and eliminates background noise while preserving detailed information presents the main challenges in this field. We propose Feature Fusion and Edge Detection Net (F2-EDNet), a camouflaged object segmentation model based on feature fusion and edge detection. ConvNeXt is used as the backbone to extract multi-scale contextual features. The extensiveness and diversity of features are then enhanced through two approaches. The first approach involves using the Feature Enhancement Module (FEM) to refine and downsize the multi-scale contextual features. The second approach introduces an auxiliary task to fuse cross-layer features through the Cross-layer Guided Edge prediction Branch (CGEB). The process extracts edge features and predicts edge information to increase feature diversity. Additionally, the Multiscale Feature Aggregation Module (MFAM) improves feature fusion by capturing and fusing information about interlayer differences between edge features and contextual features through multiscale attention and feature cascading. The model's prediction results are subjected to deep supervision to obtain the final target detection results. To validate the performance of the proposed model, it is compared qualitatively and quantitatively with eight camouflage object models from the past three years on three publicly available datasets. This comparison aims to observe its detection accuracy. Additionally, a model efficiency analysis is conducted by comparing it with five open-source models. Finally, the module's effectiveness is verified through ablation experiments to determine the optimal structure. The results of a quantitative experiment indicate that on the CAMO dataset, the S-measure, Fmeasure, E-measure correlation and mean absolute error metrics for F2-EDNet are optimal. On the COD10K dataset, the structural similarity metric indicates that the proposed algorithm is optimal, while the mean precision and recall, E-measure and MAEmetrics reach sub-optimal levels. On NC4K, all four metrics for the proposed algorithm reach optimization. From the visualized detection results, it can be observed that in the camouflage object detection task, the prediction results of the proposed model are more accurate and refined than those of other methods. Compared with other models, although the number of parameters in the proposed model is higher, the simple structure of the model framework enables it to outperform models specifically designed for lightweight purposes, faster than most other models. In comparison of the number of operations, the arithmetic complexity of the proposed model shows a significant decrease compared to a model that also utilizes multi-task learning. The model presented maintains high accuracy in target detection performance while ensuring a reasonable balance between computing speed and the number of operations. The results of ablation experiments demonstrate that each of the current modules plays the expected role, and the model's performance has been optimized. Experimental results show that the proposed algorithm achieves optimal detection accuracy. Compared to suboptimal models, our model demonstrates an average improvement of 1.41%, 1.74%, 0.14%, and 0.77% on the S- measure, F- measure, MAE, and E- measure indices across three datasets.  
Additionally, the model's design achieves a reasonable balance between operation volume and operation rate. During performance testing, the model's test speed was 46 fps, striking a balance between detection accuracy and execution efficiency, demonstrating practical application value. In future work, the algorithms will be lightened to further reduce the amount of computation to improve the speed of model inference; in applications, the model can be helpful in directions such as medical segmentation, defect detection with transparent object segmentation through migration learning.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1004-4213  
入藏号:  
WOS:001347674700003  
Indexed Date:  
2024-11-08  
Institution Address:  
Beijing Informat Sci & Technol Univ, Sch Instrumentat Sci & Optoelect Engn, Beijing 100192, Peoples R China  
Institution:

### 226. Buildings extraction of GF -2 remote sensing image based on multi -

摘要: The task of extracting buildings with high - resolution remote sensing image plays an important role in urban planning and urbanization. In view of the problems of existing deep learning extraction methods,for example,the shallow features cant been used effectively and small target information is easily lost,this paper proposes a multi - level perceptual network. This network uses dense connection mechanism to fully extract feature information,and constructs parallel structure to retain spatial information of different feature resolution and enhance feature information of different depth and scale in order to reduce the loss of detail feature. At the same time,the ASPP module is used to obtain the information of different receptive fields and extract the deep architectural features at different scales. The experimental results show that the overall accuracy of the proposed method is 97.19%, intersection over union is 74.33% and the F1 score is 85.43% in the buildings extraction of GF - 2 remote sensing image,all of which are higher than those of the traditional method and other deep learning methods. In addition, buildings with multi - source remote sensing images still have good extraction effect,which reflects the practicability of the method presented in this paper.  
摘要:  
高分辨率遥感影像建筑物提取任务在城市规划、城镇化进程等领域发挥着重要作用。针对现有的深度学习提取方法存在浅层特征未得到有效利用、小目标信息容易丢失等问题,提出了一种多层次感知网络。该网络利用密集连接机制充分提取特征信息,并构建平行结构保留不同特征分辨率的空间信息,增强不同深度、尺度特征信息,减少细节特征的丢失;同时利用空洞空间金字塔模块获取不同感受野信息,提取不同尺度下的深层建筑特征。实验结果表明,该方法在GF - 2遥感影像建筑物提取中,总体精度为97.19%、交并比为74.33%、综合评价指标为85.43%,各指标均高于传统方法与其他深度学习方法;此外,应对多源遥感影像的建筑物仍具有良好的提取效果,体现了本文方法的实用性。  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
4  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1001-070X  
入藏号:  
CSCD:6990813  
Indexed Date:  
2021-09-10  
Institution Address:  
安徽大学资源与环境工程学院, 合肥, 安徽 230601, 中国  
Institution Address:  
武汉大学, 测绘遥感信息工程国家重点实验室, 武汉, 湖北 430072, 中国  
Institution Address:  
安徽大学资源与环境工程学院;;安徽省地理信息智能技术工程研究中心, ;;安徽省地理信息智能技术工程研究中心, 合肥;;合肥, ;; 230601;;230601  
Institution Address:  
School of Resources and Environmental Engineering,Anhui University, Hefei, Anhui 230601, China  
Institution Address:  
Mapping and Remote Sensing, State Key Laboratory of Information Engineering in Surveying, Wuhan, Hubei 430072, China  
Institution Address:  
School of Resources and Environmental Engineering,Anhui University;;Anhui Engineering Research Center for Geographical Information Intelligent Technology, ;;Anhui Engineering Research Center for Geographical Information Intelligent Technology, Hefei;;Hefei, ;; 230601;;230601  
Institution:  
安徽大学资源与环境工程学院  
Institution:  
武汉大学  
Institution:  
安徽大学资源与环境工程学院;;安徽省地理信息智能技术工程研究中心  
Institution:  
School of Resources and Environmental Engineering,Anhui University  
Institution:  
Mapping and Remote Sensing  
Institution:  
School of Resources and Environmental Engineering,Anhui University;;Anhui Engineering Research Center for Geographical Information Intelligent Technology

### 227. MFSFNet: Multi-Scale Feature Subtraction Fusion Network for Remote

摘要: Change detection plays a crucial role in remote sensing by identifying surface modifications between two sets of temporal remote sensing images. Recent advancements in deep learning techniques have yielded significant achievements in this field. However, there are still some challenges: (1) Existing change feature fusion methods often introduce redundant information. (2) The complexity of network structures leads to a large number of parameters and difficulties in model training. To overcome these challenges, this paper proposes a Multi-Scale Feature Subtraction Fusion Network (MFSF-Net). It comprises two primary modules: the Multi-scale Feature Subtraction Fusion (MFSF) module and the Feature Deep Supervision (FDS) module. MFSF enhances change features and reduces redundant pseudo-change features. FDS provides additional supervision on different scales of change features in the decoder, improving the training efficiency performance of the network. Additionally, to address the problem of imbalanced samples, the Dice loss strategy is introduced as a means to mitigate this issue. Through comprehensive experiments, MFSF-Net achieves an F1 score of 91.15% and 95.64% on LEVIR-CD and CDD benchmark datasets, respectively, outperforming six state-of-the-art algorithms. Moreover, it attains an improved balance between model complexity and performance, showcasing the efficacy of the proposed approach.  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
eISSN:  
2072-4292  
入藏号:  
WOS:001045586700001  
Indexed Date:  
2023-08-19  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Chinese Acad Sci, Key Lab Technol Geo Spatial Informat Proc & Applic, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 101408, Peoples R China  
Institution:  
Institution:  
Institution:

### 228. An autonomous fault detection system for robot circuit boards based on

摘要: During the manufacturing and use of robots,it is common for robots to malfunction due to circuit failure.Therefore,the fault detection of printed circuit boards is the key to ensuring the reliability of the robot.The circuit board fault defect detection method based on deep learning can not only effectively overcome the shortcomings of manual visual inspection and online instrument detection,but also save a lot of cost for production.With the continuous development of deep learning in the field of target detection,deep neural networks can quickly and accurately identify targets from images,and have strong robustness and transferability.Aiming at the problem of circuit board fault detection,this paper proposes a robot circuit board autonomous fault detection system based on deep learning.This system greatly improves the model's detection recall rate for small faults through attention network and multiscale feature fusion,and by combining the cascaded neural network improves the accuracy of fault location.The accuracy of this model on the circuit board fault detection test data set reaches AP(IoU =0.5)= 99.4%,which fulfills the requirements of circuit board defect detection.  
摘要:  
机器人在生产制造以及使用过程中,由电路板问题导致机器人故障的情况屡见不鲜,因此印刷电路板的故障检测是保障机器人可靠性的关键。基于深度学习的电路板故障缺陷检测方法不仅能够有效克服人工目视检测和线上仪器检测的弊端,提高效率,同时能为生产节省很大的成本。随着深度学习在目标检测领域的不断发展,深度学习网络可以快速准确地从图像中识别目标,并且具有较强的鲁棒性和可迁移性。本文针对电路板的故障检测问题,提出了基于深度学习的机器人电路板自主故障检测系统,该系统通过注意力网络以及多尺度特征融合,大大提高了模型对细小故障的检测召回率,且结合了级联神经网络,提升了故障的定位精度。该模型在电路板故障检测测试数据集上的精度达到平均准确率为99.4 %,满足电路板缺陷检测的实际检测需求。  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1003-3076  
入藏号:  
CSCD:7181784  
Indexed Date:  
2022-06-10  
Institution Address:  
广东省惠州供电局, 惠州, 广东 516003, 中国  
Institution Address:  
西安交通大学电气工程学院, 西安, 陕西 710049, 中国  
Institution Address:  
Huizhou Power Supply Bureau, Huizhou, Guangdong 516003, China  
Institution Address:  
School of Electrical Engineering,Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China  
Institution:  
广东省惠州供电局  
Institution:  
西安交通大学电气工程学院  
Institution:  
Huizhou Power Supply Bureau  
Institution:  
School of Electrical Engineering,Xi'an Jiaotong University

### 229. MGRCFusion: An infrared and visible image fusion network based on

摘要: The purpose of fusing infrared and visible images is to obtain an informative image that contains bright thermal targets and rich visible texture details. However, the existing deep learning-based algorithms generally neglect finer deep-level multi-scale features, and only the last layer of features is injected into the feature fusion strategy. To this end, we propose an optimized network model for deeper-level multi-scale features extraction based on multi-scale group residual convolution. Meanwhile, a dense connection module is designed to adequately integrate these multi-scale feature information. We contrast our method with advanced deep learning-based algorithms on multiple datasets. Extensive qualitative and quantitative experiments reveal that our method surpasses the existing fusion methods. Furthermore, ablation experiments illustrate the excellence of the multi-scale group residual convolution module for infrared and visible image fusion.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0030-3992  
eISSN:  
1879-2545  
入藏号:  
WOS:001295934100001  
Indexed Date:  
2024-08-26  
Institution Address:  
Wuhan Univ Sci & Technol, Sch Machinery & Automat, Wuhan 430081, Peoples R China  
Institution:

### 230. Image Inpainting Based on Structural Constraint and Multi-Scale Feature

摘要: When repairing masked images based on deep learning, there is usually insufficient representation of multi-level information and inadequate utilization of long distance features. To solve the problems, this paper proposes a second-order generative image inpainting model based on Structural Constraints and Multi-scale Feature Fusion (SCMFF). The SCMFF model consists of two parts: edge repair network and image inpainting network. The edge repair network combines the auto-encoder with the Dilated Residual Feature Pyramid Fusion (DRFPF) module, which improves the representation of multi-level semantic information and structural details of images, thus achieves better edge repair. Then, the image inpainting network embeds the Dilated Multi-scale Attention Fusion (DMAF) module in the auto-encoder for texture synthesis with the real edge as the prior condition, and achieves fine-grained inpainting under the edge constraint by aggregating the long-distance features of different dimensions. Finally, the edge repair results are used to replace the real edge, and the two networks are fused and trained to achieve end-to-end repair from the masked image to the complete image. The model is compared with the advanced methods on datasets including Celeba, Facade and Places2. The quantitative results show that the four metrics of LPIPS, MAE, PSNR and SSIM are improved by 0.0124-0.0211, 3.787-6.829, 2.934dB-5.730dB and 0.034-0.132, respectively. The qualitative results show that the edge distribution in the center of the hole reconstructed by the SCMFF model is more uniform, and the texture synthesis effect is more in line with human visual perception.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
ISSN:  
2169-3536  
入藏号:  
WOS:000939936800001  
Indexed Date:  
2023-03-16  
Institution Address:  
Xizang Minzu Univ, Coll Informat Engn, Xianyang 712082, Shaanxi, Peoples R China  
Institution Address:  
Zhengzhou Inst Sci & Technol, Coll Informat Engn, Zhengzhou 450052, Peoples R China  
Institution Address:  
Univ CAPF, Coll Equipment Management & Guarantee, Xian 712000, Peoples R China  
Institution:  
Institution:  
Zhengzhou Inst Sci & Technol  
Institution:  
Univ CAPF

### 231. Defect detection in automotive glass based on modified YOLOv5 with

摘要: Automotive glass is one of the key components in manufacturing engineering, and the inspection of defects is essentially an important item of quality evaluation. Deep learning has become a promising technology and is very suitable for defect detection. However, a unified method to detect all kinds of defects is very difficult due to the few publicly available image datasets of automotive glass, and it is a challenge to deploy high-precision models in resource-constrained edge devices. Focusing on these problems, defect detection in automotive glass based on modified YOLOv5 with pseudo-tagging and dual lightweight strategy is developed in this paper. First, aiming at the problem that the defect samples are extremely lacking, combined pseudo-labeling and traditional data expansion methods have been explored to effectively increase the number of samples for meeting the requirement of deep learning model training, thus improving the performance of the detection model. Second, double lightweight modules, MobileNetV3, and Ghost module are introduced into the backbone and the neck network of the YOLOv5 model, respectively, for reducing the complexity of the model. In addition, a multi-scale feature fusion (M-SFF) module, which riches the semantic and spatial information of feature maps, is added to the output of the backbone to further improve the detection accuracy of the model. The effectiveness of each innovation module was verified through ablation experiment and horizontal comparative experimentally. Experimental results show that the complexity and deployment difficulty of the improved YOLOv5 model are significantly reduced, the floating-point operations per second (FLOPs) values and weight sizes are significantly smaller than those of other models, and the detection performance is also improved based on the original YOLOv5, an average accuracy increase of 3.9% on average, precision slightly decreased, and recall rates increased by 7.6%. The method is appropriate for applications in the defect inspection of automotive glass.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001143010300001  
Indexed Date:  
2024-01-31  
Institution Address:  
Fujian Key Lab Automot Elect & Elect Drive, Fuzhou 350118, Peoples R China  
Institution Address:  
Fujian Univ Technol, Sch Elect Elect Engn & Phys, Fuzhou 350118, Peoples R China  
Institution:  
Fujian Key Lab Automot Elect & Elect Drive  
Institution:

### 232. Hybrid supervised metal surface defect detection based on multi-scale

摘要: Aiming at the problem in defect detection that effective features cannot be extracted due to different shapes of defect targets in the detected samples, this paper presents a defect detection model based on deep learning, which uses an improved multi-scale feature fusion module to solve the problem of identifying defects of different sizes on the basis of controlling the amount of calculation. By introducing a non-local attention mechanism module, the models ability of extracting defect features is enhanced. Furthermore, mixed-supervised training is used in training to explore the relationship between the amount of annotations required by the model and the detection accuracy. This method achieves better accuracy than the state-of-the-art methods on KSDD, KSDD2, and STEEL datasets, and can extract discriminative features for different types of defects. Compared with the state-of-the-art fully supervised and unsupervised methods, the average accuracy improvement on the dataset is 0.8% and 11%.  
摘要:  
针对缺陷检测中被检测样品中因缺陷目标形状各异引起的无法提取有效特征的问题,本文提出基于深度学习的缺陷检测模型。该模型使用改进后的多尺度特征融合模块,在控制计算量的基础上解决识别不同大小缺陷的问题。通过引入非局部注意力机制模块,模型对缺陷特征的提取能力得到加强;在训练中使用混合监督训练,探索模型所需要的标注量和检测准确度之间的关系。本文方法在KSDD、KSDD2和STEEL 3个数据集上都获得了比先进方法更好的精确度,对于不同类型的缺陷都能提取到有判别力的特征。与先进的完全监督方法和无监督方法相比,在数据集上精确度平均提高0.8%和11%。  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1673-4785  
入藏号:  
CSCD:7531475  
Indexed Date:  
2023-11-24  
Institution Address:  
常州大学计算机与人工智能学院、阿里云大数据学院、软件学院, 常州, 江苏 213000, 中国  
Institution Address:  
School of Compute Science and Artificial Intelligence / Aliyun School of Big Data / School of Software, Changzhou University, Changzhou, 213000  
Institution:  
常州大学计算机与人工智能学院、阿里云大数据学院、软件学院  
Institution:  
School of Compute Science and Artificial Intelligence / Aliyun School of Big Data / School of Software, Changzhou University

### 233. Multi-Scale Feature Fusion Network with Symmetric Attention for Land

摘要: The complementary characteristics of SAR and optical images are beneficial in improving the accuracy of land cover classification. Deep learning-based models have achieved some notable results. However, how to effectively extract and fuse the unique features of multi-modal images for pixel-level classification remains challenging. In this article, a two-branch supervised semantic segmentation framework without any pretrained backbone is proposed. Specifically, a novel symmetric attention module is designed with improved strip pooling. The multiple long receptive fields can better perceive irregular objects and obtain more anisotropic contextual information. Meanwhile, to solve the semantic absence and inconsistency of different modalities, we construct a multi-scale fusion module, which is composed of atrous spatial pyramid pooling, varisized convolutions and skip connections. A joint loss function is introduced to constrain the backpropagation and reduce the impact of class imbalance. Validation experiments were implemented on the DFC2020 and WHU-OPT-SAR datasets. The proposed model achieved the best quantitative values on the metrics of OA, Kappa and mIoU, and its class accuracy was also excellent. It is worth mentioning that the number of parameters and the computational complexity of the method are relatively low. The adaptability of the model was verified on RGB-thermal segmentation task.  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
eISSN:  
2072-4292  
入藏号:  
WOS:001192991600001  
Indexed Date:  
2024-04-04  
Institution Address:  
Chinese Acad Sci, Changchun Inst Opt Fine Mech & Phys, Changchun 130033, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 234. Accurate prediction of ice surface and bottom boundary based on

摘要: Identifying the locations of ice surface and bottom boundary in the radar imagery enables the calculation of ice sheet thickness, which is one of important inputs for ice-sheet modelling and global climate research. Therefore, accurate predictions of the boundaries can contribute to improve the accuracy of global climate analysis and sea level prediction. However, an accurate boundary detection in radar sounder data collected from the polar ice sheet has still been a challenge because the boundaries of the ice layer are usually very weak and noisy, and subglacial topography is highly variable. In recent years, the deep learning methods have surpassed the performances of traditional technology and helped to overcome a series of problems, including image boundary segmentation and target detection. This paper proposes a multi-scale feature fusion network (MFFN) for boundary detection of ice sheet radar echograms, where the ground truth supervises the output of the network at different stages, rather than the output of the last layer of the network. Also, a multi-scale convolution module (MCM) is introduced to learn the rich multi-scale representation of each network stage from shallow to deep, which uses convolution with different dilation rates to obtain multi-scale features. Furthermore, an improved loss function makes the proposed MFFN more effective to solve the sample imbalance problem of boundary detection, and further improves the accuracy of boundary detection. The proposed method is verified experimentally using the radar echograms from 2009 provided by the Center of Remote Sensing of Ice Sheets (CReSIS) that are used as training and test data. In the experiments, the proposed MFFN not only achieves state-of-the-art boundary detection accuracy on the test set but also improves the visual effect by generating fine boundaries.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000772273400001  
Indexed Date:  
2022-03-31  
Institution Address:  
Beijing Univ Technol, Fac Informat Technol, Beijing 100124, Peoples R China  
Institution:

### 235. A Novel Hybridoma Cell Segmentation Method Based on Multi-Scale Feature

摘要: The hybridoma cell screening method is usually done manually by human eyes during the production process for monoclonal antibody drugs. This traditional screening method has certain limitations, such as low efficiency and subjectivity bias. Furthermore, most of the existing deep learning-based image segmentation methods have certain drawbacks, due to different shapes of hybridoma cells and uneven location distribution. In this paper, we propose a deep hybridoma cell image segmentation method based on residual and attention U-Net (RA-UNet). Firstly, the feature maps of the five modules in the network encoder are used for multi-scale feature fusion in a feature pyramid form and then spliced into the network decoder to enrich the semantic level of the feature maps in the decoder. Secondly, a dual attention mechanism module based on global and channel attention mechanisms is presented. The global attention mechanism (non-local neural network) is connected to the network decoder to expand the receptive field of the feature map and bring more rich information to the network. Then, the channel attention mechanism SENet (the squeeze-and-excitation network) is connected to the non-local attention mechanism. Consequently, the important features are enhanced by the learning of the feature channel weights, and the secondary features are suppressed, hence improving the cell segmentation performance and accuracy. Finally, the focal loss function is used to guide the network to learn the hard-to-classify cell categories. Furthermore, we evaluate the performance of the proposed RA-UNet method on a newly established hybridoma cell image dataset. Experimental results show that the proposed method has good reliability and improves the efficiency of hybridoma cell segmentation compared with state-of-the-art networks such as FCN, UNet, and UNet++. The results show that the proposed RA-UNet model has improvements of 0.8937%, 0.9926%, 0.9512%, and 0.9007% in terms of the dice coefficients, PA, MPA, and MIoU, respectively.  
Web of Science 核心合集中的 "被引频次":  
32  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
32  
eISSN:  
2079-9292  
入藏号:  
WOS:000945012200001  
Indexed Date:  
2023-03-23  
Institution Address:  
Hangzhou Dianzi Univ, Sch Comp Sci & Technol, Hangzhou 310018, Peoples R China  
Institution Address:  
Zhejiang Police Coll, Minist Publ Secur, Key Lab Publ Secur Informat Applicat Based Big Dat, Hangzhou 310000, Peoples R China  
Institution Address:  
Menoufia Univ, Fac Artificial Intelligence, Shibin Al Kawm 32511, Egypt  
Institution:  
Institution:  
Institution:

### 236. A multi-scale feature fusion network based fast CU partitioning in HEVC

摘要: High Efficiency Video Coding (HEVC) significantly improves the coding efficiency but increases the coding complexity,especially in the process of coding unit (CU) partitioning based on quadtree structure,so it is important to study the fast CU partitioning.A multi-scale feature fusion network can achieve fast HEVC CU partitioning.Therefore,the UcuNet network structure is designed by combining the U-Net and CU partitioning features.Meanwhile,asymmetric convolutional AC and CBAM attention mechanisms are used to enhance the feature extraction of pixels at different scales.In order to sufficiently train the deep learning model,the original video with different resolutions and the corresponding encoding information are collected to build a large-scale dataset.Finally,the model is embedded into the HEVC coding architecture to predict the result of CU partitioning in advance,which can effectively reduce the coding complexity caused by CU partitioning by eliminating the recursive rate distortion optimization (RDO) calculation process in the original CU partitioning method.Compared with the official HEVC test model (HM16.20),the proposed UcuNet reduces the average coding time by 68.13% while BD-BR is only decreased by 2.63%.  
摘要:  
高效视频编码HEVC显著提高了编码效率,但同时增加了编码复杂度,在基于四叉树结构的编码单元(CU)划分过程中尤为明显,因此研究CU快速划分具有重要意义。多尺度特征融合的网络可以实现HEVC编码单元快速划分。为此,结合U-Net和CU划分特性设计了UcuNet网络,同时为加强不同尺度像素的特征提取,采用了非对称卷积AC和CBAM注意力机制。为更好地训练深度学习模型,收集了不同分辨率的原始视频和对应的编码信息构建出大规模的数据集。最后将模型嵌入到HEVC编码架构中,提前预测CU划分的结果,跳过了原始CU划分方法中递归的率失真优化(RDO)计算过程,从而有效降低CU划分带来的编码复杂度。实验结果表明,对比HEVC官方测试模型(HM16.20),UcuNet在BD-BR仅损失2.63%的情况下,使平均编码时间缩短了68.13%。  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
1  
ISSN:  
1007-130X  
入藏号:  
CSCD:7606963  
Indexed Date:  
2024-02-02  
Institution Address:  
河北工业大学电子信息工程学院;;河北工业大学电子与通信工程国家级实验教学示范中心, ;;, ;;, 天津;;天津 300131;;300131, 中国  
Institution Address:  
广东工业大学信息工程学院, 广州, 广东 510006, 中国  
Institution Address:  
天津大学电气自动化与信息工程学院, 天津 300072, 中国  
Institution Address:  
School of Electronic and Information Engineering,Hebei University of Technology;;National Experimental Teaching Demonstration Center of Electronic and Communication Engineering,Hebei University of Technology, ;;, ;;, Tianjin;;Tianjin 300131;;300131  
Institution Address:  
School of Information Engineering,Guangdong University of Technology, Guangzhou, Guangdong 510006, China  
Institution Address:  
School of Electrical and Information Engineering,Tianjin University, Tianjin 300072, China  
Institution:  
河北工业大学电子信息工程学院;;河北工业大学电子与通信工程国家级实验教学示范中心  
Institution:  
广东工业大学信息工程学院  
Institution:  
天津大学电气自动化与信息工程学院  
Institution:  
School of Electronic and Information Engineering,Hebei University of Technology;;National Experimental Teaching Demonstration Center of Electronic and Communication Engineering,Hebei University of Technology  
Institution:  
School of Information Engineering,Guangdong University of Technology  
Institution:  
School of Electrical and Information Engineering,Tianjin University

### 237. Double multi-scale feature fusion network for crowd counting

摘要: Recently, the research of crowd counting has attracted increasing attention but still faces many challenges, such as crowded scenes, scale variations and cluttered backgrounds. With the development of deep learning, density maps are widely used for crowd counting, where the quality of density maps plays a crucial role in counting performance. In this paper, we propose a new convolutional network architecture, called double multi-scale feature fusion network (DMFFNet), to generate high-quality density maps and accurate counting estimates. DMFFNet utilizes VGG19 to extract multi-scale feature maps from input images. The features from last three scales are further enlarged the receptive fields by three designed dilated feature pyramid modules, and then fused together. Moreover, a feature enhancement module composed of spatial attention and channel-wise attention is presented to weight the fused feature maps for effectively distinguishing between crowd and background. We also design a new dual-scale loss to optimize the network during training. Experimental results show that DMFFNet reduces MAEs by at least 1.5%\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\%$$\end{document}, 1.5%\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\%$$\end{document}, 1.2%\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\%$$\end{document}, 0.6%\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\%$$\end{document} and 0.5%\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\%$$\end{document} on UCF\_\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\\_$$\end{document}CC\_\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\\_$$\end{document}50, UCF-QNRF, JHU-Crowd++, ShanghaiTech Part A and Part B datasets, and decreases MSEs by at least 1.8%\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\%$$\end{document} and 0.1%\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\%$$\end{document} on JHU-Crowd++ and ShanghaiTech Part B datasets, as compared with the state-of-the-art.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001176735100016  
Indexed Date:  
2024-03-28  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Artificial Intelligence, Sch Future Technol, Nanjing 210044, Peoples R China  
Institution:

### 238. AUTOMATIC DETECTION AND GRADING OF DIABETIC MACULAR EDEMA BASED ON A

摘要: Purpose: To solve the problem of automatic grading of macular edema in retinal images in a more stable and reliable way and reduce the workload of ophthalmologists, an automatic detection and grading method of diabetic macular edema based on a deep neural network is proposed.  
Methods: The enhanced green channels of fundus images are input into the YOLO network for training and testing. Diabetic macular edema is graded according to the distance of the macula and hard exudate. We used multiscale feature fusion to form more comprehensive features on different grain images to improve the effect of hard exudate detection. We adopted K-means++ algorithm to cluster anchor box size and use LGIoU loss of the original network to guide the regression of hard exudate bounding box and improve the regression accuracy of anchor boxes. We increased the diversity of samples for sample training by data augmentation, including cropping, flipping, and rotating of fundus images, so that each batch of training data can better represent the distribution of samples.  
Results: The detection accuracy of the proposed method can reach 96% on the MESSIDOR data set. The detection rates of hard exudate with high, median, and low probability are 100%, 79.12%, and 60.40%, respectively.  
Conclusion: The proposed method exhibits a very good detection stability on healthy and diseased fundus images.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
0275-004X  
eISSN:  
1539-2864  
入藏号:  
WOS:000798904200014  
PubMed ID:  
35152245  
Indexed Date:  
2022-06-01  
Institution Address:  
Jilin Univ, Key Lab Symbol Computat & Knowledge Engn, Minist Educ, Changchun, Jilin, Peoples R China  
Institution Address:  
Jilin Univ, Coll Comp Sci & Technol, 2699 Qianjin St, Changchun 130012, Jilin, Peoples R China  
Institution Address:  
Bethune First Hosp Jilin Univ, Ophthalmol Dept, Changchun, Jilin, Peoples R China  
Institution Address:  
Bethune Second Hosp Jilin Univ, Ophthalmol Dept, Changchun, Jilin, Peoples R China  
Institution:  
Institution:  
Institution:  
Bethune First Hosp Jilin Univ  
Institution:  
Bethune Second Hosp Jilin Univ

### 239. End-to-End Learnable Multi-Scale Feature Compression for VCM

摘要: The proliferation of deep learning-based machine vision applications has given rise to a new type of compression, so called video coding for machine (VCM). VCM differs from traditional video coding in that it is optimized for machine vision performance instead of human visual quality. In the feature compression track of MPEG-VCM, multi-scale features extracted from images are subject to compression. Recent feature compression works have demonstrated that the versatile video coding (VVC) standard-based approach can achieve a BD-rate reduction of up to 96% against MPEG-VCM feature anchor. However, it is still sub-optimal as VVC was not designed for extracted features but for natural images. Moreover, the high encoding complexity of VVC makes it difficult to design a lightweight encoder without sacrificing performance. To address these challenges, we propose a novel multi-scale feature compression method that enables both the end-to-end optimization on the extracted features and the design of lightweight encoders. The proposed model combines a learnable compressor with a multi-scale feature fusion network so that the redundancy in the multi-scale features is effectively removed. Instead of simply cascading the fusion network and the compression network, we integrate the fusion and encoding processes in an interleaved way. Our model first encodes a larger-scale feature to obtain a latent representation and then fuses the latent with a smaller-scale feature. This process is successively performed until the smallest-scale feature is fused and then the encoded latent at the final stage is entropy-coded for transmission. The results show that our model outperforms previous approaches by at least 52% BD-rate reduction and has x5 to x27 times less encoding time for object detection. It is noteworthy that our model can attain near-lossless task performance with only 0.002-0.003% of the uncompressed feature data size.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
被引频次合计:  
6  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:001221132000073  
Indexed Date:  
2024-06-15  
Institution Address:  
Kyung Hee Univ, Sch Comp, Visual Media Lab, Yongin 17104, South Korea  
Institution Address:  
Elect & Telecommun Res Inst ETRI, Daejeon 34129, South Korea  
Institution:  
Institution:

### 240. Multi-scale triple-attention network for pixelwise crack segmentation

摘要: Currently, intelligent crack detection is of great value for the maintenance of infrastructure, of which the most significant kind in China is roads. For pavement defects, the pavement can be repaired and maintained in a timely manner with an accurate defect detection task, which significantly reduces the occurrence of hazards. However, the detection of pavement defects remains a great challenge owing to many difficulties, for example, complex backgrounds, microdefects, various defect shapes and sizes, class imbalance issues, etc. Recently, deep learning has demonstrated its superior performance on pixelwise image segmentation, but some issues still exist on demanding pixelwise image segmentation, for instance, limited receptive field, insufficiency processing of local features, information loss issue generated by pooling operations, etc. Based on all of the above issues, a multiscale triple-attention network, named MST-Net, is proposed for end-to-end pixelwise crack detection. First, a multiscale input strategy is applied to the proposed segmentation network to capture more context information. Meanwhile, it can capably reduce the effect of the information loss issue generated by pooling operations. Second, to realize effective feature representation of local features, an additive attention fusion (AAF) block is proposed to guide feature learning to capture both global and local contexts. In addition, faced with the crack detection task with class imbalance issues, a triple attention (TA) block is proposed to detect spatial, channel and pixel attention information to suppress the background and useless information, which is conducive to the characterization of microcracks. Finally, aiming at the limited receptive field, a multiscale feature aggregation unit is proposed for feature fusion to increase the detection ability of multiscale defects. To better guide network training, a deep supervision mechanism is also introduced to speed up the convergence of the proposed segmentation model and improve the performance of defect segmentation. The related evaluation and detection experiments are carried out on three public datasets on crack segmentation, and the comparison experiments with the mainstream segmentation models show that the proposed segmentation network achieves excellent performance on pixelwise crack detection.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
39  
被引频次合计:  
40  
ISSN:  
0926-5805  
eISSN:  
1872-7891  
入藏号:  
WOS:000975127500001  
Indexed Date:  
2023-05-09  
Institution Address:  
Zhengzhou Univ, Sch Elect & Informat Engn, Henan 450001, Peoples R China  
Institution Address:  
Robot Percept & Control Engn Lab Henan Prov, Zhengzhou 450001, Henan, Peoples R China  
Institution Address:  
Edinburgh Napier Univ, Built Environm, Edinburgh EH10 5DT, Scotland  
Institution:  
Institution:  
Robot Percept & Control Engn Lab Henan Prov  
Institution:

### 241. Method of oil and gas reservoir detection based on geological knowledge

摘要: Oil and gas reservoir detection is one of the major tasks of petroleum energy companies in the exploration and production process. The oil and gas industry has long relied on the expert manual analysis of massive logging data to perform qualitative analyses of oil and gas reservoirs. Although experts interpretations are highly accurate, the time and economic costs are considerably high. With the rapid development of artificial intelligence technologies such as deep learning in recent years, intelligent oil and gas reservoir detection methods have become a focus in the academia and industry. However, sensor data in real industrial scenarios present serious inconsistencies, which bring great challenges to traditional supervised learning models. This paper presents a focused study on the oil and gas reservoir detection task in the context of sensor inconsistencies and proposes a geological knowledge distillation multiscale network approach. This method proposes a multiscale feature fusion mechanism based on self-attention to learn the multiscale dynamic representation of geological information. Then, the model designs a geological knowledge distillation learning framework to learn additional geological knowledge from inconsistent sensor data. This step further improves the models accuracy. A large number of experiments on real industrial datasets are subsequently performed. The results fully prove the effectiveness and robustness of the proposed model in oil and gas reservoir detection.  
摘要:  
油气储集层识别是石油能源企业在勘测和开发业务中核心的任务之一.长期以来,油气行业一直依靠专家人工分析海量测井数据以对地下油气储集层进行定性分析,虽然专家解释结论有着很高的精准度,但是时间与经济成本都十分高昂.近些年来,随着以深度学习为代表的人工智能技术的迅速发展,智能油气储集层识别技术成为学术界和工业界共同关注的问题.然而,真实工业环境存在严重的传感数据不一致问题,给传统的监督学习模型带来巨大的挑战.本文针对传感器不一致情境中油气储集层识别任务展开研究,提出多尺度地质知识蒸馏网络的方法.首先,该方法提出一种多尺度特征自注意力融合机制来学习地质信息的多尺度动态表征.其次,该方法设计一种地质知识蒸馏学习模型,从非一致传感数据中学习额外的地质知识,进一步提升模型准确度.最后,在真实数据集上进行大量实验,结果充分证明本文提出的模型在油气储集层识别任务上的有效性和鲁棒性.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1674-7267  
入藏号:  
CSCD:6920946  
Indexed Date:  
2021-06-04  
Institution Address:  
中国科学技术大学, 大数据分析与应用安徽省重点实验室, 合肥, 安徽 230027, 中国  
Institution Address:  
华为技术有限公司, 杭州, 浙江 310051, 中国  
Institution Address:  
中国石油勘探开发研究院, 北京 100083, 中国  
Institution Address:  
University of Science and Technology of China, Anhui Province Key Laboratory of Big Data Analysis and Applicationn, Hefei, Anhui 230027, China  
Institution Address:  
Huawei Technologies Co., Ltd., Hangzhou, Zhejiang 310051, China  
Institution Address:  
Research Institute of Petroleum Exploration and Development, Beijing 100083, China  
Institution:  
中国科学技术大学  
Institution:  
华为技术有限公司  
Institution:  
中国石油勘探开发研究院  
Institution:  
University of Science and Technology of China  
Institution:  
Huawei Technologies Co., Ltd.  
Institution:  
Research Institute of Petroleum Exploration and Development

### 242. Bidirectional Multi-scale Deformable Attention for Video

摘要: Video super-resolution aims to generate a high-resolution video frame from its low-resolution video sequences. Video super-resolution is still a challenging problem due to performing the temporal frame alignment and spatial feature fusion during the process of spatial-temporal modeling. Existing deep learning based methods have limitations in handling accurate alignment and effective fusion of frames with multi-scale feature information. In this paper, we propose Bidirectional Multi-scale Deformable Attention (BMDA) for video Super-Resolution in terms of propagation, alignment and fusion. More specifically, the developed Deformable Alignment Module (DAM) in BMDA contains two kinds of modules: Multi-scale Deformable Convolution Module (MDCM) and Multi-scale Attention Module (MAM). MDCM is leveraged to deal with the offset information in different scales and align adjacent frames at the feature level, improving the robustness of the alignment among adjacent frames. MAM is designed to extract the local and global features of the aligned features for aggregation, such that the feature information compensation between pixels is achieved. Additionally, in order to make full use of shallow features, dense connection structure between each layer is adopted in the framework of bidirectional propagation to achieve better visual performance on video super-resolution. In particular, our proposed BDAM outperforms BasicVSR by up to 1.28dB in PSNR when batch size is set to 2. Experimental results on public video benchmark datasets demonstrate that the proposed method can achieve superior performance on large motion videos as compared with the state-of-the art methods.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001051861800003  
Indexed Date:  
2023-09-01  
Institution Address:  
Zhejiang Univ Finance & Econ, Sch Data Sci, Hangzhou 310018, Peoples R China  
Institution Address:  
China Jiliang Univ, Coll Sci, Dept Data Sci, Hangzhou 310018, Zhejiang, Peoples R China  
Institution Address:  
Murdoch Univ, Discipline Engn & Energy, Perth, WA 6150, Australia  
Institution Address:  
China Jiliang Univ, Coll Informat Engn, Hangzhou 310018, Zhejiang, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 243. BSNet: a boundary-aware medical image segmentation network

摘要: Accurate segmentation of medical images can provide foundations for clinical and disease diagnosis. Inaccurate segmentation boundaries often result from limited contextual information and insufficient discriminating feature maps after consecutive pooling and upsampling operations in most existing methods. In this paper, we present a novel boundary-aware medical image segmentation network (BSNet) for resolving the multi-objective segmentation problem. We exploit a backbone network to extract multi-scale feature representations and design an adaptive contrast boundary-aware module (ACB), which uses the method of combining nonlinear filters with deep learning to extract high-quality boundary maps. We then build a feature fusion (FF) module to fuse multi-scale features with boundary maps, providing decoder with rich multi-scale features enhanced with boundary information, and facilitating cross-channel interactions. To further enhance the uncertain regions of the boundaries, we utilize the boundary spatial enhancement (BSE) module to learn the feature map of boundary locations with the assistance of the Sobel operator. We conducted experiments with three challenging public datasets to evaluate the effectiveness of BSNet. Simulation results on various datasets show that the present model outperforms state-of-the-art segmentation methods, obtaining up to 2.73% improvement in Dice coefficient (DICE) score. BSNet opens new ways of designing better boundary-aware segmentation network.Please confirm the corresponding author is correctly identified.No problem.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2190-5444  
入藏号:  
WOS:001402539400004  
Indexed Date:  
2025-02-01  
Institution Address:  
Jiaotong Univ, Sch Informat Sci & Technol, Chengdu 610031, Peoples R China  
Institution Address:  
Dublin City Univ, Sch Elect Engn, Dublin, Ireland  
Institution:  
Jiaotong Univ  
Institution:

### 244. AsymUNet: An Efficient Multi-Layer Perceptron Model Based on Asymmetric

摘要: With the continuous advancement of deep learning technology, U-Net-based algorithms for image denoising play a crucial role in medical image processing. However, most U-Net-based medical image denoising algorithms typically have large parameter sizes, which poses significant limitations in practical applications where computational resources are limited or large-scale patient data processing are required. In this paper, we propose a medical image denoising algorithm called AsymUNet, developed using an asymmetric U-Net framework and a spatially rearranged multilayer perceptron (MLP). AsymUNet utilizes an asymmetric U-Net to reduce the computational burden, while a multiscale feature fusion module enhances the feature interaction between the encoder and decoder. To better preserve the image details, spatially rearranged MLP blocks serve as the core building blocks of AsymUNet. These blocks effectively extract both the local and global features of the image, reducing the model's reliance on prior knowledge of the image and further accelerating the training and inference processes. Experimental results demonstrate that AsymUNet achieves superior performance metrics and visual results compared with other state-of-the-art methods.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2079-9292  
入藏号:  
WOS:001305991500001  
Indexed Date:  
2024-09-15  
Institution Address:  
Nanjing Normal Univ Special Educ, Sch Math & Informat Sci, Nanjing 210038, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Software, Nanjing 210044, Peoples R China  
Institution Address:  
Inner Mongolia Univ Technol, Sch Informat Engn, Hohhot 010051, Peoples R China  
Institution Address:  
Nanjing Univ, State Key Lab Novel Software Technol, Nanjing 210023, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Jiangsu Prov Engn Res Ctr Adv Comp & Intelligent S, Nanjing 210044, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 245. Redundant object detection method for civil aircraft assembly based on

摘要: Slow detection of redundant objects and low accuracy in assembly lines, particularly in the setting of civil aircraft assembly, are tough and challenging problems. To address these issues, a redundant object detection method based on computer vision and augmented reality (AR) smart glasses is proposed in this paper. The method uses AR glasses as the image collection hardware and takes the live image collected by the camera as the input of the proposed deep learning machine vision model. The proposed model, the Feature Pyramid Networks-CenterNet, is inspired by CenterNet and combined with multi-scale feature fusion to solve the problem of low detection accuracy of small-scale redundant targets. The weight factor of the loss function was set according to the proportion of small targets in the dataset, which solves the problem of an unbalanced proportion of large and small targets in the training samples. The proposed network model was validated on the PASCAL Visual Object Classes public dataset and the self-built redundant object dataset. The results showed that the new method can detect seven redundant objects with a mean accuracy of 74.49% within the visible range of smart glasses within 200 ms. The research provides a new reference for the quality process management of civil aircraft assembly.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
10  
被引频次合计:  
12  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:000832751800001  
Indexed Date:  
2022-08-05  
Institution Address:  
Shanghai Univ, Sch Mechatron Engn & Automat, Shanghai, Peoples R China  
Institution:

### 246. Accelerating Cross-Scene Co-Seismic Landslide Detection Through

摘要: Sudden co-seismic landslides strike, causing widespread devastation and demanding a rapid response. The swift and accurate acquisition of landslide information is essential for effective disaster relief. Deep learning (DL)-based computer-aided interpretation methods have emerged as cutting-edge tools for landslide detection. Nevertheless, traditional DL approaches face limitations, such as high annotation costs, slow processing speeds, and low generalizability, rendering them unsuitable for rapid co-seismic landslide recognition tasks. This study presents a progressive approach for co-seismic landslide detection. First, we develop a Multi-scale Feature Fusion Lightweight Neural Network (MFFLnet), achieving exceptional generalizability and speed while maintaining precision. Second, we employ the deep transfer learning (TL) strategy, enabling MFFLnet to leverage prior landslide knowledge from a source domain and a refined data augmentation algorithm to combat overfitting. The proposed methodology is implemented in two co-seismic landslide scenes in Hokkaido, Japan, and Luding, China. Experimental results demonstrate that the proposed method exhibits outstanding performance in regional landslide recognition and robust performance across different co-seismic landslide detection scenarios. Our approach proves competitive in efficient co-seismic landslide disaster recognition and cross-scene identification, showcasing significant applicability in the face of rapid response demands.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
9  
被引频次合计:  
9  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001272260000017  
Indexed Date:  
2024-09-06  
Institution Address:  
China Univ Geosci, Badong Natl Observat & Res Stn Geohazards, Wuhan 430074, Peoples R China  
Institution Address:  
China Univ Geosci, Natl Engn Res Ctr Geog Informat Syst, Wuhan 430074, Peoples R China  
Institution Address:  
Hohai Univ, Geotech Res Inst, Nanjing 210098, Peoples R China  
Institution Address:  
Nippon Koei Co Ltd, Geohazard Management Dept, Geohazard Management Div, Tokyo 1028539, Japan  
Institution:  
Institution:  
Institution:  
Institution:

### 247. Improving small object detection via context-aware and feature-enhanced

摘要: Detecting small objects is a challenging task in computer vision due to the objects only occupying a limited number of pixels and having blurred contours. These factors result in minimal discriminative features being available to effectively model the objects. In this paper, we propose three lightweight plug-and-play modules that can be seamlessly integrated into object detection algorithms, particularly those in the YOLO series, to improve the accuracy of detecting small objects. The Spatially Enhanced Convolutional Block Attention Module (SE-CBAM) is integrated into the feature extraction layer of the network to enhance the feature extraction capability of neural networks. Additionally, a Contextual Information Pooling Enhancement Module (CIE-Pool) is included at the multi-scale feature fusion stage to extract and improve object background information, which enhances the recognition rate of small objects. To improve the detection of small objects, a new layer is added to the detection head, which incorporates the shallow feature map obtained from the feature extraction network after Adaptive Feature Processing (AFP), thereby obtaining more and richer information about small objects. The efficacy of the algorithm has been evaluated on the VisDrone2021 and AI-TOD datasets. The experimental results demonstrate that the method proposed in this paper greatly improves the detection accuracy of small objects while maintaining real-time capabilities. Furthermore, it maintains high accuracy and speed even when dealing with complex background conditions and detecting small objects with high blur.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001172780300002  
Indexed Date:  
2024-03-17  
Institution Address:  
Hangzhou Dianzi Univ, Coll Elect & Informat, Hangzhou 310018, Zhejiang, Peoples R China  
Institution Address:  
Zhejiang Dahua Technol Co Ltd, Hangzhou 310053, Zhejiang, Peoples R China  
Institution:  
Institution:  
Zhejiang Dahua Technol Co Ltd

### 248. Automatic Extraction of Urban Impervious Surface Based on SAH-Unet

摘要: Increases in the area of impervious surfaces have occurred with urbanization. Such surfaces are an important indicator of urban expansion and the natural environment. The automatic extraction of impervious surface data can provide useful information for urban and regional management and planning and can contribute to the realization of the United Nations Sustainable Development Goal 11-Sustainable Cities and Communities. This paper uses Google Earth Engine (GEE) high-resolution remote sensing images and OpenStreetMap (OSM) data for Chengdu, a typical city in China, to establish an impervious surface dataset for deep learning. To improve the extraction accuracy, the Small Attention Hybrid Unet (SAH-Unet) model is proposed. It is based on the Unet architecture but with attention modules and a multi-scale feature fusion mechanism. Finally, depthwise-separable convolutions are used to reduce the number of model parameters. The results show that, compared with other classical semantic segmentation networks, the SAH-Unet network has superior precision and accuracy. The final scores on the test set were as follows: Accuracy = 0.9159, MIOU = 0.8467, F-score = 0.9117, Recall = 0.9199, Precision = 0.9042. This study provides support for urban sustainable development by improving the extraction of impervious surface information from remote sensing images.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
8  
eISSN:  
2072-4292  
入藏号:  
WOS:000942278900001  
Indexed Date:  
2023-03-23  
Institution Address:  
Chengdu Univ Technol, Coll Math & Phys, Chengdu 610059, Peoples R China  
Institution Address:  
Chengdu Univ Technol, Digital Hu Line Res Inst, Chengdu 610059, Peoples R China  
Institution Address:  
Chengdu Univ Technol, Geomath Key Lab Sichuan Prov, Chengdu 610059, Peoples R China  
Institution Address:  
Int Res Ctr Big Data Sustainable Dev Goals CBAS, Beijing 100094, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Key Lab Digital Earth Sci, Beijing 100094, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 249. Res2-UNet++: a deep learning image post-processing method for electrical

摘要: The monitoring of multiphase flow distribution in industrial processes in order to optimize production presents a challenge. Electrical resistance tomography (ERT) is a technique used to visualize the inner distribution of multiphase flow. Image reconstruction plays a vital role in ERT. However, the nonlinearity and ill-posedness of inverse problems make image reconstruction in ERT difficult. The development of advanced imaging algorithms has attracted much interest for this purpose. In this work, an improved U-shaped deep learning model is proposed, which combines the advantages of the multi-scale feature extraction of UNet++ and the residual feature fusion of Res2Net. The network is designed to post-process the pre-reconstruction results of traditional ERT image-reconstruction methods, combining the generalization ability of the model-based methods and the flexible feature-extraction advantage of deep learning methods. The post-processing includes super-resolution, image denoising and artifact removal. Simulations and experiments are designed to verify the generalization ability and effectiveness of the proposed post-processing model. Both simulation and experimental results show that the proposed U-shaped network approach outperforms other deep learning methods, and the proposed deep learning model is fit for post-processing in ERT, making it a robust solution for applications.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001268874500001  
Indexed Date:  
2024-07-22  
Institution Address:  
Tianjin Univ, Sch Elect & Informat Engn, Tianjin Key Lab Proc Measurement & Control, Tianjin 300072, Peoples R China  
Institution Address:  
Tianjin Renai Coll, Sch Informat & Intelligent Engn, Tianjin, Peoples R China  
Institution:  
Institution:  
Tianjin Renai Coll

### 250. DMFNet: A Novel Self-Supervised Dynamic Multi-Focusing Network for

摘要: In recent years, speech denoising has greatly benefited from the rapid development of neural networks. However, these models require substantial noisy-clean speech pairs for supervised training, which limits their widespread use. Although there have been attempts to train denoising networks with only noisy speech data, existing self-supervised methods often suffer from a lack of continuity, low noise reduction performance, or heavy dependence on noise modeling. In this work, we introduce an efficient self-supervised Dynamic Multi-Focusing Network (DMFNet), a noise-only trained speech denoising network that utilizes a multi-scale connected encoder-decoder architecture as its backbone. Specifically, we have designed an efficient Spectral Dynamic Focusing Unit (SDFU) that enables the network to dynamically adapt the shape of its convolutional kernels while learning features, thus effectively focusing on the spectral structure of the human voice. Additionally, we introduce a Complex Attention Module (CAM), designed with a cross-space structure specialized for feature interaction and extraction. Finally, to further enhance the recovery of fine spectral details, we propose the Complex Multi-Scale Feature Fusion Unit (CMFFU) and Complex Scope Fusion Unit (CSFU) to adaptively fuse the features from different stages in the encoding process. Extensive evaluations across multiple datasets demonstrate that the proposed DMFNet significantly outperforms other state-of-the-art methods.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001276350900001  
Indexed Date:  
2024-08-01  
Institution Address:  
Xiamen Univ, Dept Marine Technol & Engn, Xiamen 361005, Peoples R China  
Institution Address:  
Beijing Univ Chem Technol, Coll Informat Sci & Technol, Beijing 100029, Peoples R China  
Institution:  
Institution:

### 251. SKS-Net: skeleton-strengthening network for surface defect segmentation

摘要: Surface defect detection is an important task in industrial production. Although significant progress has been made in deep learning-based defect segmentation methods, the low contrast between defect and background and the shape and scale diversity of defects limit the models' detection accuracy and generalization ability. Therefore, realizing the full automation of surface defect detection still faces many challenges. To overcome these problems, this paper proposes a skeleton-strengthening network called SKS-Net, which provides stable and excellent surface defect detection performance even at low contrast and multi-scale. SKS-Net designs a skeleton-strengthening convolutional module to capture multi-scale features efficiently. The convolution kernel of this convolution module is closer to the shape of the segmentation target, which significantly reduces irrelevant regions and improves the feature extraction capability of the convolution kernel. This convolution module can be directly embedded into existing network structures without adding additional computational overhead. In addition, we design a new feature channel fusion module to extract key information from features at different levels. To improve the training effect, we introduce a multi-scale auxiliary supervision mechanism. The proposed model is evaluated on four different publicly available surface defect datasets and compared with other state-of-the-art models. Results show that SKS-Net performs exceptionally well in terms of accuracy, achieving 66.72% mIoU on the KolektorSDD dataset. The code is publicly available at https://github.com/Wanglaoban3/SKS-Net.git.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001388370500001  
Indexed Date:  
2025-01-07  
Institution Address:  
China Univ Petr East China, Coll New Energy, Qingdao 266580, Shandong, Peoples R China  
Institution Address:  
Hunan Inst Adv Technol, Ultraprecis Mfg Collaborat Innovat Ctr, Changsha 410072, Hunan, Peoples R China  
Institution:  
Institution:  
Hunan Inst Adv Technol

### 252. A self-interpretable deep learning network for early prediction of

摘要: Accurate prediction of pathologic complete response to neoadjuvant chemotherapy non-invasively before treatment via dynamic contrast-enhanced magnetic resonance imaging is vital for developing a personalized therapy strategy. However, the application of deep learning in this domain is characterized by its black-box nature, largely relying on post-hoc analysis to interpret final decision-making. This reliance results in a lack of self-interpretability in the operational mechanisms of feature extraction, feature fusion, and decision-making. Moreover, these models have demonstrated unsatisfactory prediction performance due to insufficient feature modeling. To address these issues, we propose a self-interpretable deep learning network that can provide the intrinsic interpretability of feature extraction, multi-scale feature fusion, and final prediction. First, the interpretable perception module is designed to extract features both effectively and interpretably. Furthermore, the interpretable adaptive multi-scale feature fusion module is proposed to fuse multi-scale features. Finally, an endto-end self-interpretable deep learning network is presented to predict pathologic complete response with selfinterpretability. Validated on a multi-center pre-treatment dynamic contrast-enhanced magnetic resonance imaging dataset, our self-interpretable deep learning network outperforms state-of-the-art methods in both prediction performance and self-interpretability, improving the area under the receiver operating characteristic curve by at least 4.81% while providing both qualitative and quantitative self-interpretability. Our study demonstrates that our proposed self-interpretable deep learning network can extract key information from pretreatment breast dynamic contrast-enhanced magnetic resonance imaging while enhancing both the prediction performance and the transparency of the model, thereby improving its trustworthiness in clinical settings.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001335100100001  
Indexed Date:  
2024-10-23  
Institution Address:  
Univ Sci & Technol Beijing, Sch Automat & Elect Engn, Beijing 100083, Peoples R China  
Institution Address:  
Minist Educ, Key Lab Knowledge Automat Ind Proc, Beijing 100083, Peoples R China  
Institution:  
Institution:  
Minist Educ

### 253. YOLO-UAV: Object Detection Method of Unmanned Aerial Vehicle Imagery

摘要: As Unmanned Aerial Vehicle (UAV) remote sensing technology progresses, the utilization of deep learning in UAV imagery object detection has become more prevalent. However, detecting small targets in complex backgrounds and distinguishing dense targets remains a major challenge. To address these issues and improve object detection efficiency, this study proposes an UAV imagery object detection method called YOLO-UAV by optimizing YOLOv5. YOLO-UAV first reconstructs the backbone and feature fusion networks by simplifying the network structure and reducing computational burden. The employment of a Dense\_CSPDarknet53 backbone network, fashioned via the incorporation of dense connections, facilitates the extraction of latent image information through the recurrent utilization of features. In the Neck structure, an efficient feature fusion block with structural re-parameterization and ELAN strategies is integrated to effectively reduce interference from complex background noise while extracting more accurate and rich features. In addition, by proposing GS-Decoupled Head, this approach diminishes the parameter count of the decoupled head without compromising accuracy. It also separates classification tasks from regression tasks to lessen the influence of task disparities on prediction bias. To tackle the discrepancy between positive and negative samples in bounding box regression tasks, this study introduces a new loss function, Focal-ECIoU, capable of expediting network convergence and improve model positioning ability. Experimental findings from the public VisDrone2019 dataset indicate that YOLO-UAV outperforms other advanced object detection methods in comprehensive performance. Compared with the baseline model YOLOv5s, YOLO-UAV increased mAP0.5 from 35.1% to 46.7%, while mAP0.5:0.95 increased from 19.1% to 27.4%. For small-scale targets, AP(small) increased from 10.2% to 17.3%. The experiment proves that YOLO-UAV performs well in improving object detection accuracy and has strong generalization ability, satisfying the practical requirements of UAV imagery object detection tasks.  
Web of Science 核心合集中的 "被引频次":  
6  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
2169-3536  
入藏号:  
WOS:001105487100001  
Indexed Date:  
2023-12-03  
Institution Address:  
Xinjiang Univ, Sch Informat Sci & Engn, Sch Cyberspace Secur, Urumqi 830017, Peoples R China  
Institution Address:  
Beijing Acad Sci & Technol, Beijing 100035, Peoples R China  
Institution Address:  
Xinjiang Lianhaichuangzhi Informat Technol Co Ltd, Key Lab Big Data Xinjiang Social Secur Risk, Urumqi 830011, Peoples R China  
Institution:  
Institution:  
Institution:  
Xinjiang Lianhaichuangzhi Informat Technol Co Ltd

### 254. Incipient fault diagnosis of analog circuit with ensemble HKELM based on

摘要: As an essential part in electronics-rich system, the failure of analog circuits will severely affect the system reliability and security. Incipient fault of analog circuit refers to the early stage of degradation fault where the fault characteristics are generally weak and almost indistinguishable. In order to enhance the reliability of electronic systems, it is necessary to diagnose incipient faults of analog circuits promptly and effectively. Existing approaches generally capture fault characteristics only from single signal, ignoring the valuable information inherent in different domains and scales. To address this problem, a novel diagnostic strategy based on multi-scale feature extraction and multi-channel feature fusion is designed to guarantee the completeness and richness of fault information. In this study, a deep extreme learning machine denoising auto-encoder (DELM-DAE) based method is proposed to conduct unsupervised multi-scale and multi-channel feature fusion to extract distinguishable features for incipient faults. The proposed method has higher learning efficiency and overcomes the common problem of low efficiency in deep learning model training. Meanwhile, in order to improve the ability to distinguish high-resolution features, an ensemble hybrid kernel extreme learning machine with novel roulette selection and weighted voting scheme is proposed to enhance the recognition performance and stability. In the verification experiment, the diagnosis accuracy on four typical circuits all reaches above 98%, which demonstrates that the proposed incipient fault diagnosis method for analog circuits has more conspicuous performance than other state-of-the-art methods.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
18  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
20  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:000993663200001  
Indexed Date:  
2023-06-04  
Institution Address:  
Northwestern Polytech Univ, Sch Civil Aviat, Xian 710072, Peoples R China  
Institution Address:  
Xian Univ Architecture & Technol, Sch Mech & Elect Engn, Xian 710055, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, 127 Youyi West Rd, Xian 710072, Shaanxi, Peoples R China  
Institution:  
Institution:  
Institution:

### 255. Multiscale and Multilevel Feature Fusion Network for Quantitative

摘要: Passive microwave (PMW) radiometers have been widely utilized for quantitative precipitation estimation (QPE) by leveraging the relationship between brightness temperature (Tb) and rain rate. Nevertheless, accurate precipitation estimation remains a challenge due to the intricate relationship between them, which is influenced by a diverse range of complex atmospheric and surface properties. In addition, the inherent skew distribution of rainfall values prevents models from correctly addressing extreme precipitation events, leading to a significant underestimation. This article presents a novel model called the multiscale and multilevel feature fusion network (MSMLNet), consisting of two essential components: a multiscale feature extractor and a multilevel regression predictor. The feature extractor is specifically designed to extract characteristics from multiple scales, enabling the model to incorporate various meteorological conditions, as well as atmospheric and surface information in the surrounding environment. The regression predictor first assesses the probabilities of multiple rainfall levels for each observed pixel and then extracts features of different levels separately. The multilevel features are fused according to the predicted probabilities. This approach allows each submodule only to focus on a specific range of precipitation, avoiding the undesirable effects of skew distributions. To evaluate the performance of MSMLNet, various deep learning methods are adapted for the precipitation retrieval task, and a PWM-based product from the global precipitation measurement (GPM) mission is also used for comparison. Extensive experiments show that MSMLNet surpasses GMI-based products and the most advanced deep learning approaches by 17.9% and 2.5% in root mean square error (RMSE), and 54.2% and 4.0% in CSI-10, respectively. Moreover, we demonstrate that MSMLNet significantly mitigates the propensity for underestimating heavy precipitation events and has a consistent and outstanding performance in estimating precipitation across various levels.  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001226385000009  
Indexed Date:  
2024-06-16  
Institution Address:  
Harbin Inst Technol, Dept Comp Sci, Shenzhen Key Lab Internet Informat Collaborat, Shenzhen 518055, Peoples R China  
Institution Address:  
China Meteorol Adm, Natl Satellite Meteorol Ctr, Beijing 100081, Peoples R China  
Institution:  
Institution:

### 256. A generalized well neural network for surface defect segmentation in

摘要: Surface defect detection is an important task in the field of manufacturing, and dealing with imbalanced data is a challenge that has been addressed using methods such as anomaly detection and data augmentation. However, optical devices pose a particular challenge due to their characteristics of small batches and varying types, resulting in insufficient positive sample data and difficulty in predicting the data distribution of new batches. To address this issue, we propose a neural network that learns to compare the differences between templates and testing samples, rather than directly learning the representations of the samples. By collecting templates, the model can generalize to new batches. The challenge of extracting defect features by comparison is to remove background noise, such as displacements, deformations, and texture changes. We propose a Dual -Attention Mechanism (DAM) in the stage of feature extraction, which extracts the noise-free defect features using the non-position information of self-attention. In the stage of feature fusion, we introduce a Recurrent Residual Attention Mechanism (RRAM) to generate spatial masks that shield noise and enable multi-scale feature fusion. We evaluate our method on three datasets of Optical Communication Devices (OCDs), Printed Circuit Boards (PCBs) and Motor Commutator Surface Defects (MCSD), and demonstrate that it outperforms existing state-of-the-art methods. Our work provides a promising direction for addressing the challenge of surface defect detection in OCDs and can be generalized to other flexible manufacturing system (FMS).  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
10  
ISSN:  
0166-3615  
eISSN:  
1872-6194  
入藏号:  
WOS:001025163000001  
Indexed Date:  
2023-07-24  
Institution Address:  
Huazhong Univ Sci & Technol, Sch Mech Sci & Engn, 1037 Luoyu Rd, Wuhan 430074, Hubei, Peoples R China  
Institution Address:  
Wuhan Intelligent Equipment Ind Inst Co Ltd, 8 Ligou South Rd, Wuhan 430074, Hubei, Peoples R China  
Institution Address:  
Natl Univ Singapore, Coll Design & Engn, Dept Mech Engn, Singapore 117575, Singapore  
Institution:  
Institution:  
Wuhan Intelligent Equipment Ind Inst Co Ltd  
Institution:

### 257. Fourier ptychography based on multi-scale feature fusion network

摘要: Fourier Ptychography (FP) is a technology of achieving high-resolution, large field-of-view imaging of optical system. However, the high-resolution reconstruction based on traditional FP methods requires a high aperture overlap ratio, resulting in a large number of captured images and low sampling efficiency. In addition, the FP reconstruction algorithm has high complexity and long reconstruction time. Aiming at solving these problems of the FP, this paper proposes a deep learning algorithm based on multi-scale feature fusion network. Through the improved feature pyramid module, the feature information can be extracted from multiple low-resolution images captured by the FP imaging system, and the information is fused to achieve super-resolution reconstruction. Experimental results show that compared with traditional methods, the deep learning algorithm proposed in this paper improves the quality of image reconstruction, reduces the reconstruction time by 90%, and is more robust to Gaussian noise. In addition, the proposed method can reduce the overlap ratio between sub-apertures from 50% to 25% in frequency domain, and reduce the number of captured images by 50%, greatly improving the sampling efficiency.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1007-2780  
eISSN:  
2097-3217  
入藏号:  
WOS:000883684200011  
Indexed Date:  
2022-11-21  
Institution Address:  
Chinese Acad Sci, Changchun Inst Opt Mech & Phys, Changchun 130033, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution Address:  
Lappeenranta Lahti Univ Technol, Sch Engn Sci, Comp Vis & Pattern Recognit Lab, Lahti 15210, Finland  
Institution:  
Institution:  
Institution:

### 258. Camouflage Object Detection Based on Feature Fusion and Edge Detection

摘要: Camouflaged Object Detection (COD) holds significant research and application value in various fields. The ability of deep learning is pushing the performance of target detection algorithms to new heights. Designing a network that effectively integrates features of different layer sizes and eliminates background noise while preserving detailed information presents the main challenges in this field. We propose Feature Fusion and Edge Detection Net (F2-EDNet), a camouflaged object segmentation model based on feature fusion and edge detection. ConvNeXt is used as the backbone to extract multi-scale contextual features. The extensiveness and diversity of features are then enhanced through two approaches. The first approach involves using the Feature Enhancement Module (FEM) to refine and downsize the multi-scale contextual features. The second approach introduces an auxiliary task to fuse cross-layer features through the Cross-layer Guided Edge prediction Branch (CGEB). The process extracts edge features and predicts edge information to increase feature diversity. Additionally, the Multiscale Feature Aggregation Module (MFAM) improves feature fusion by capturing and fusing information about interlayer differences between edge features and contextual features through multiscale attention and feature cascading. The model's prediction results are subjected to deep supervision to obtain the final target detection results. To validate the performance of the proposed model, it is compared qualitatively and quantitatively with eight camouflage object models from the past three years on three publicly available datasets. This comparison aims to observe its detection accuracy. Additionally, a model efficiency analysis is conducted by comparing it with five open-source models. Finally, the module's effectiveness is verified through ablation experiments to determine the optimal structure. The results of a quantitative experiment indicate that on the CAMO dataset, the S-measure, Fmeasure, E-measure correlation and mean absolute error metrics for F2-EDNet are optimal. On the COD10K dataset, the structural similarity metric indicates that the proposed algorithm is optimal, while the mean precision and recall, E-measure and MAEmetrics reach sub-optimal levels. On NC4K, all four metrics for the proposed algorithm reach optimization. From the visualized detection results, it can be observed that in the camouflage object detection task, the prediction results of the proposed model are more accurate and refined than those of other methods. Compared with other models, although the number of parameters in the proposed model is higher, the simple structure of the model framework enables it to outperform models specifically designed for lightweight purposes, faster than most other models. In comparison of the number of operations, the arithmetic complexity of the proposed model shows a significant decrease compared to a model that also utilizes multi-task learning. The model presented maintains high accuracy in target detection performance while ensuring a reasonable balance between computing speed and the number of operations. The results of ablation experiments demonstrate that each of the current modules plays the expected role, and the model's performance has been optimized. Experimental results show that the proposed algorithm achieves optimal detection accuracy. Compared to suboptimal models, our model demonstrates an average improvement of 1.41%, 1.74%, 0.14%, and 0.77% on the S-measure, F-measure, MAE, and E-measure indices across three datasets. Additionally, the model's design achieves a reasonable balance between operation volume and operation rate. During performance testing, the model's test speed was 46 fps, striking a balance between detection accuracy and execution efficiency, demonstrating practical application value. In future work, the algorithms will be lightened to further reduce the amount of computation to improve the speed of model inference; in applications, the model can be helpful in directions such as medical segmentation, defect detection with transparent object segmentation through migration learning.  
摘要:  
针对伪装目标边缘模糊、相关检测模型上下文特征利用率低、边缘特征融合繁琐的问题,提出一种基于特征融合与边缘检测的伪装目标检测模型F2-EDNet。首先构造特征增强模块,细化主干网络的多尺度上下文特征,有效增强伪装目标特征信息;同时,引入跨层特征引导的边缘预测支路以集成来自主干网络底层和顶层的跨层特征,在辅助检测伪装目标边缘的同时,提取边缘特征;最后,提出多尺度特征聚合模块,通过结合注意力机制,充分融合边缘特征与上下文特征,有效提高预测精度。实验结果表明,F2-EDNet在公开数据集CAMO、COD10K和NC4K上的结构相似性、平均精度与召回率、相关性、平均误差指标均值分别提高了1.41%、1.74%、0.14%、0.77%;和同类模型相比,该模型具有更丰富的边缘,定位伪装区域更准确;在实际应用中,模型检测速率可达46帧/s,证明模型具有较好的实时检测能力。  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1004-4213  
入藏号:  
CSCD:7806939  
Indexed Date:  
2024-11-23  
Institution Address:  
北京信息科技大学仪器科学与光电工程学院, 北京 100192, 中国  
Institution Address:  
School of Instrumentation Science and Opto-electronics Engineering,Beijing Information Science and Technology University, Beijing 100192, China  
Institution:  
北京信息科技大学仪器科学与光电工程学院  
Institution:  
School of Instrumentation Science and Opto-electronics Engineering,Beijing Information Science and Technology University

### 259. CAMU-Net: Copy-move forgery detection utilizing coordinate attention and

摘要: In this paper, we construct CAMU-Net, an image forgery detection method, to obtain evidence of copy-move forgery areas in images. In CAMU-Net, the hierarchical feature extraction stage (HFE\_Stage) is used to extract multi-scale key feature maps. Next, a hierarchical feature matching stage (HFM\_Stage) based on self-correlation combined with a multi-scale structure is designed to predict copy-move forgery areas with different scales of information. To optimize the matching results, we design a coordinate attention-based resource allocation stage (CARA\_Stage), which uses a location and channel attention mechanism to assign more weight to copy-move areas. In this way, useful information can be strengthened while irrelevant information is suppressed. To effectively use the multi-scale prediction results in the multi-scale feature fusion-based up-sampling stage (MFFU\_Stage), we integrate the high-level and low-level information into one information flow. By combining the global feature information of the deep layers and the location details of the shallow layers, the performance of CMFD can be improved. To demonstrate the validity of our model, we compare it with a variety of traditional methods and deep learning methods. The results show that our performance is outstanding. In particular, on the COVERAGE dataset, our AUC is 87.3%, which is 2.4% higher than the second place. In addition, we design a variety of baseline methods to perform several ablation experiments to demonstrate the validity of the modules in this model.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001097352500001  
Indexed Date:  
2023-11-18  
Institution Address:  
Macau Univ Sci & Technol, Sch Comp Sci & Engn, Macau 999078, Peoples R China  
Institution Address:  
Macao Polytech Univ, Fac Appl Sci, R Luis Gonzaga Gomes, Macau 999078, Peoples R China  
Institution Address:  
Guangdong Univ Technol, Sch Comp Sci & Technol, Guangzhou 510000, Peoples R China  
Institution:  
Institution:  
Institution:

### 260. 被撤回的出版物: Evaluation of the Impact of the Introduction of Overseas

摘要: For the central cities, the transformation and upgrading of the industrial structure, the continuous deepening of innovation capabilities, and the improvement of sustainable development capabilities, human resources are playing an increasingly important role. In the modern era, assembling an army of highly educated, high-quality, and highly talented personnel has become a practical demand of talent development strategy. In order to respond to national policies, keep up with the trend of the times, and give play to the comparative advantages of talent competition, the primary task of central cities is to do a good job in the introduction of overseas high-level talents. It is necessary for the central cities to establish a mechanism for the introduction of overseas high-level talents with the government as an important task and to formulate targeted talent introduction policies according to the local characteristics of the central cities. It is very important to evaluate the impact of the introduction of overseas high-level talents on the development of central cities so that the central cities can formulate talent introduction strategies according to the actual situation. This work uses an artificial neural network to evaluate the impact of the introduction of overseas high-level talents on the development of central cities. Aiming at the problem that the evaluation accuracy and computing efficiency of the deep learning-based method decrease due to the proliferation of neural network layers, and the improved residual network model is proposed. On the one hand, a multiscale feature fusion block is designed in the first layer of the network model, which can extract the multiscale feature information in the original special. On the other hand, the residual block is optimized and improved by using depthwise separable convolution to remove the computational burden on the network.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
4  
ISSN:  
1939-0114  
eISSN:  
1939-0122  
入藏号:  
WOS:000805157000014  
Indexed Date:  
2022-06-13  
Institution Address:  
Jingchu Univ Technol, Foreign Language Sch, Jingmen 448000, Hubei, Peoples R China  
Institution Address:  
Natl Inst Dev Adm, Int Coll, Bangkok 10240, Thailand  
Institution:  
Institution:

### 261. DCTE-LLIE: A Dual Color-and-Texture-Enhancement-Based Method for

摘要: The enhancement of images captured under low-light conditions plays a vitally important role in the area of image processing and can significantly affect the performance of following operations. In recent years, deep learning techniques have been leveraged in the area of low-light image enhancement tasks, and deep-learning-based low-light image enhancement methods have been the mainstream for low-light image enhancement tasks. However, due to the inability of existing methods to effectively maintain the color distribution of the original input image and to effectively handle feature descriptions at different scales, the final enhanced image exhibits color distortion and local blurring phenomena. So, in this paper, a novel dual color-and-texture-enhancement-based low-light image enhancement method is proposed, which can effectively enhance low-light images. Firstly, a novel color enhancement block is leveraged to help maintain color distribution during the enhancement process, which can further eliminate the color distortion effect; after that, an attention-based multiscale texture enhancement block is proposed to help the network focus on multiscale local regions and extract more reliable texture representations automatically, and a fusion strategy is leveraged to fuse the multiscale feature representations automatically and finally generate the enhanced reflection component. The experimental results on public datasets and real-world low-light images established the effectiveness of the proposed method on low-light image enhancement tasks.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2073-431X  
入藏号:  
WOS:001254571900001  
Indexed Date:  
2024-07-02  
Institution Address:  
Chinese Acad Sci, Xian Inst Opt & Precis Mech, Xian 710119, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 262. Ship Detection in SAR Images Based on Multi-Scale Feature Extraction and

摘要: Deep learning has attracted increasing attention across a number of disciplines in recent years. In the field of remote sensing, ship detection based on deep learning for synthetic aperture radar (SAR) imagery is replacing traditional methods as a mainstream research method. The multiple scales of ship objects make the detection of ship targets a challenging task in SAR images. This paper proposes a new methodology for better detection of multi-scale ship objects in SAR images, which is based on YOLOv5 with a small model size (YOLOv5s), namely the multi-scale ship detection network (MSSDNet). We construct two modules in MSSDNet: the CSPMRes2 (Cross Stage Partial network with Modified Res2Net) module for improving feature representation capability and the FC-FPN (Feature Pyramid Network with Fusion Coefficients) module for fusing feature maps adaptively. Firstly, the CSPMRes2 module introduces modified Res2Net (MRes2) with a coordinate attention module (CAM) for multi-scale features extraction in scale dimension, then the CSPMRes2 module will be used as a basic module in the depth dimension of the MSSDNet backbone. Thus, our backbone of MSSDNet has the capabilities of features extraction in both depth and scale dimensions. In the FC-FPN module, we set a learnable fusion coefficient for each feature map participating in fusion, which helps the FC-FPN module choose the best features to fuse for multi-scale objects detection tasks. After the feature fusion, we pass the output through the CSPMRes2 module for better feature representation. The performance evaluation for this study is conducted using an RTX2080Ti GPU, and two different datasets: SSDD and SARShip are used. These experiments on SSDD and SARShip datasets confirm that MSSDNet leads to superior multi-scale ship detection compared with the state-of-the-art methods. Moreover, in comparisons of network model size and inference time, our MSSDNet also has huge advantages with related methods.  
Web of Science 核心合集中的 "被引频次":  
61  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
被引频次合计:  
65  
eISSN:  
2072-4292  
入藏号:  
WOS:000754586300001  
Indexed Date:  
2022-03-02  
Institution Address:  
Xidian Univ, Sch Aerosp Sci & Technol, Xian 710126, Peoples R China  
Institution Address:  
Shaanxi Acad Aerosp Technol Applict Co Ltd, Xian 710199, Peoples R China  
Institution:  
Institution:  
Shaanxi Acad Aerosp Technol Applict Co Ltd

### 263. Multiscale Attention-UNet-Based Near-Real-Time Precipitation Estimation

摘要: Extreme precipitation events greatly threaten people's daily lives and safety, making accurate and timely precipitation estimation especially critical. However, common methods like radar and satellite remote sensing have limitations due to coverage and environmental factors. Existing deep learning models struggle with complex scenarios and multisource data correlations. These make the precipitation estimation tasks challenging. This article proposes a Multiscale Dual Cross-Attention UNet (MS-DCA-UNet) model for near-real-time precipitation estimation. It integrates Doppler weather radar and FY-4A satellite data to overcome single-source data limitations. To narrow the semantic gap among the encoder feature maps, the MS-DCA-UNet model introduces a dual-cross attention (DCA) module at the skip connections of the backbone network U-Net. The DCA module mainly employs a channel cross-attention and a spatial cross-attention to capture remote dependencies and enable multiscale feature fusion. A multiscale convolution module is designed to reduce the risk of the model falling into local optima. It is a multibranch upsampling strategy that runs parallel to the decoder. Experimental results show that the Critical Success Index (CSI), Root Mean Square Error (RMSE), and Pearson's Correlation Coefficient (CC) of MS-DCA-UNet are 0.6033, 0.5949 mm/h, and 0.8460, respectively, with the hourly CMPAS precipitation as the benchmark. These outperform the other comparisons, such as FY-4A QPE, GPM IMERG, U-Net, Attention-UNet, and DCA-UNet on the CSI, RMSE, and CC metrics. MS-DCA-UNet reduces the RMSE of Attention-UNet, UNet, and DCA-UNet by a margin of 34.68% (0.5949 mm/h versus 0.9107 mm/h), 10.24% (0.5949 mm/h versus 0.6628 mm/h), 6.96% (0.5949 mm/h versus 0.6394 mm/h), respectively.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001360472900005  
Indexed Date:  
2024-11-30  
Institution Address:  
Chengdu Univ Informat & Technol, Sch Comp Sci, Chengdu 610000, Peoples R China  
Institution:

### 264. Edge Intelligent Perception Method for Power Grid Icing Condition Based

摘要: Insulator is an important equipment of power transmission line. Insulator icing can seriously affect the stable operation of power transmission line. So insulator icing condition monitoring has great significance of the safety and stability of power system. Therefore, this paper proposes a lightweight intelligent recognition method of insulator icing thickness for front-end ice monitoring device. In this method, the residual network (ResNet) and feature pyramid network (FPN) are fused to construct a multi-scale feature extraction network framework, so that the shallow features and deep features are fused to reduce the information loss and improve the target detection accuracy. Then, the full convolution neural network (FCN) is used to classify and regress the iced insulator, so as to realize the high-precision identification of icing thickness. Finally, the proposed method is compressed by model quantization to reduce the size and parameters of the model for adapting the icing monitoring terminal with limited computing resources, and the performance of the method is verified and compared with other classical method on the edge intelligent chip.  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
2296-598X  
入藏号:  
WOS:000717195300001  
Indexed Date:  
2021-11-21  
Institution Address:  
Wuhan Univ, Sch Elect & Automat, Wuhan, Peoples R China  
Institution Address:  
Wuhan Text Univ, Sch Comp & Artificial Intelligence, Wuhan, Peoples R China  
Institution Address:  
State Grid Xi Power Supply Co, State Grid Shaanxi Elect Power Co, Xian, Peoples R China  
Institution Address:  
Huizhou Power Supply Bur Co Ltd, Guangdong Power Grid Corp, Huizhou, Peoples R China  
Institution Address:  
Qinghai Univ, Tus Inst Renewable Energy, Xining, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 265. Multiscale common-private feature adversarial decoupling network for

摘要: Hyperspectral pansharpening involves the fusion of a high-resolution panchromatic (PAN) image with a lower- resolution hyperspectral (LRHS) image, yielding a remarkable high-resolution hyperspectral (HRHS) image. Most of the existing pansharpening methods design specific feature extraction modules to dig deep spatial- spectral features. However, these methods ignore the commonalities and characteristics between PAN image and HS image, which may cause common information redundancy and private information loss. Here, we propose a multiscale common-private feature decoupling network based on adversarial learning (called HPMFDN) for HS pansharpening, which refines and integrates the common-private features extracted from PAN and LRHS images of different scales losslessly, enhancing the pansharpening performance by fully utilizing the complementary information of PAN and HS images. Specifically, in each scale of PAN and HS images, the co-learning common-private feature decoupling module (CL-CPFDM) consisting of adversarial learning network and the specific decoupling losses is presented to decouple PAN and HS features into mutually orthogonal and independent common-private features. In addition, we specially design an information lossless refinement- based fusion module (ILRFM) for private information integration based on invertible neural network (INN), ensuring an effective spatial-spectral information flow for HRHS image reconstruction. Experimental results demonstrate that the proposed HP-MFADN outperforms other widely accepted state-of-the-art methods in both objective metrics and visual appearance. The code link is: https://github.com/Jiahuiqu/HP-MFDN.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0950-7051  
eISSN:  
1872-7409  
入藏号:  
WOS:001407419900001  
Indexed Date:  
2025-02-03  
Institution Address:  
Xidian Univ, State Key Lab Integrated Serv Network, Xian 710071, Peoples R China  
Institution Address:  
Beijing Elect Sci & Technol Inst, Dept Elect & Commun Engn, Beijing 100070, Peoples R China  
Institution Address:  
Xidian Univ, Sch Telecommun Engn, Xian 710071, Peoples R China  
Institution:  
Institution:  
Institution:

### 266. PhysKANNet: A KAN-based model for multiscale feature extraction and

摘要: Physiological indicator reflects the health status of the human body, and remote photoplethysmography (rPPG) is a highly promising technology for contactless measurement of these indicators through facial video. However, current deep learning methods mainly rely on traditional neural networks with limited spatiotemporal receptive fields, overlooking the importance of multi-scale features and noise resistance in rPPG signal modeling. This results in challenges when addressing subtle color changes and noise interference. To overcome these limitations, we leverage the advantages of the Kolmogorov-Arnold Network (KAN) in handling sparse data and propose PhysKANNet, a novel KAN-based encoder-decoder architecture that integrates multi-scale feature extraction and contextual information fusion to enhance rPPG signal extraction. We introduce three new plug-and-play modules for PhysKANNet: the rPPG-Aware Convolutional Attention Block, which extracts features at different scales through a multi-branch structure and enhances multi-scale representation using KAN's nonlinear modeling capabilities; the Multi-Dimensional Feature Fusion Block, which combines high-dimensional features from the encoder with low-dimensional features from the decoder; and the rPPG Edge Sampling Block, which fuses edge and semantic information to further optimize signal extraction accuracy. We employ unsupervised learning for training PhysKANNet and conducted comprehensive experiments on multiple benchmark datasets. The results show that PhysKANNet significantly improves feature learning from unlabeled data, achieving excellent performance across various testing scenarios.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:001349805700001  
Indexed Date:  
2024-11-15  
Institution Address:  
Chongqing Univ Technol, Sch Artificial Intelligence, Chongqing 401135, Peoples R China  
Institution Address:  
Chongqing Normal Univ, Sch Comp & Informat Sci, Chongqing 401331, Peoples R China  
Institution:  
Institution:

### 267. Earthquake signal detection using a multiscale feature fusion network

摘要: Signal and noise classification can add an extra level of constraint for earthquake phase picking by pinpointing the signal waveforms from continuous seismic data for more accurate arrival picking. However, the continuously increasing data collected by worldwide stations exceeds the ability of manual analysis. Moreover, manual earthquake data analysis depends on seismologists' expert knowledge, resulting in inconsistent analysis results. To address this, we proposed a generalized deep learning (DL) network architecture to discriminate earthquake signal and noise waveforms. The proposed DL framework is a novel architecture comprising a feature extractor, a classifier and two hybrid attention modules. It utilizes different kernel sizes for more detailed feature extraction, and the hybrid attention mechanism module can guide the network to focus more on the waveform characteristics. To illustrate the power of the proposed DL network, we applied it to classify the earthquake signal and noise of the 3-C Texas Earthquake Dataset. The results demonstrate that the accuracy of the proposed method in the testing set reaches 99.83 per cent. We further utilize the transfer learning strategy to demonstrate the transferability of the proposed network with the Stanford earthquake data set, showing an encouraging classification accuracy of 95.03 per cent. Additionally, we conducted an additional experiment on arrival picking by integrating decoder blocks into the classification network, which achieves remarkable P- and S-wave arrival picking accuracy.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0956-540X  
eISSN:  
1365-246X  
入藏号:  
WOS:001379487000001  
Indexed Date:  
2024-12-22  
Institution Address:  
Yangtze Univ, Coll Geophys & Petr Resources, Minist Educ, Key Lab Explorat Technol Oil & Gas Resources, Wuhan 430100, Peoples R China  
Institution Address:  
Univ Texas Austin, John A & Katherine G Jackson Sch Geosci, Bur Econ Geol, Univ Stn Box X, Austin, TX 78713 USA  
Institution:  
Institution:

### 268. DFSDNet: A dual-branch multi-scale feature fusion network for surface

摘要: Surface defect detection is a research hotspot in the field of computer vision. Due to the complex characteristics of metal surfaces and the multitude of industrial defects, it remains a challenging task. In order to meet the requirement of accurate identification of surface defects on copper strips and plates in industrial quality control, we propose a computer vision-based dual-branch features fusion neural network named as DFSDNet. We gather defect samples to construct the KUST-DET dataset of surface defects on copper strips and plates to support the training and evaluation of detection models. Experiments on the KUST-DET dataset demonstrate that DFSDNet-s achieved a mean Average Precision (mAP) of 88.53%, while maintaining low computational complexity and low parameters, and the model achieves a good balance between detection precision and computational efficiency. In addition, the mAP on the NEU-DET dataset is 75.67%, showcasing its good defect detection performance. Experiments have shown that DFSDNet is an effective surface defect detection model with great potential in other metal industry applications.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0166-3615  
eISSN:  
1872-6194  
入藏号:  
WOS:001433540600001  
Indexed Date:  
2025-03-06  
Institution Address:  
Kunming Univ Sci & Technol, Fac Informat Engn & Automat, Kunming 650500, Peoples R China  
Institution:

### 269. Homogeneous teacher based buffer knowledge distillation for tiny neural

摘要: Knowledge Distillation (KD) has shown great promise in improving the performance of tiny neural networks. Most existing KD methods have the large teacher-student discrepancy, thus, students hardly learn useful knowledge and may not achieve effective distillation. In this paper, we focus on the construction and training of homogeneous teachers and propose the novel Buffer Knowledge Distillation (BKD), which reduces the teacher-student discrepancy in terms of network architecture and distilled knowledge. Particularly, we first construct a series of homogeneous networks with larger capacity based on the student. A width-by-width fine-tuning mechanism is developed to reduce training costs of homogeneous networks, and the one with the highest accuracy is selected as the teacher. Furthermore, we propose BKD to reduce the learning difficulty, in which teacher and student features are fused into buffer features by our new multi-scale feature fusion module. Extensive experiments for image classification have been conducted to verify the homogeneous teacher based BKD, which consistently outperforms many existing KD methods. The results show that our method achieves up to 4.75% accuracy improvement on CIFAR-100, and the width-by-width fine-tuning mechanism incurs 33.58% and 36.08% less training time on CIFAR-100 and CIFAR-10, respectively.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1383-7621  
eISSN:  
1873-6165  
入藏号:  
WOS:001185090300001  
Indexed Date:  
2024-04-10  
Institution Address:  
East China Normal Univ, Dept Comp Sci & Technol, Shanghai, Peoples R China  
Institution Address:  
China FAW Technol Dev Co Ltd, Nanjing, Peoples R China  
Institution:  
Institution:  
China FAW Technol Dev Co Ltd

### 270. Advancing Image Object Detection: Enhanced Feature Pyramid Network and

摘要: In the era of artificial intelligence, the significance of images and videos as intuitive conveyors of information cannot be overstated. Computer vision techniques rooted in deep learning have revolutionized our ability to autonomously and accurately identify objects within visual media, making them a focal point of contemporary research. This study addresses the pivotal role of image object detection, particularly in the contexts of autonomous driving and security surveillance, by presenting an in-depth exploration of this field with a focus on enhancing the feature pyramid network. One of the key challenges in existing object detection methodologies lies in mitigating information loss caused by multi-scale feature fusion. To tackle this issue, we propose the enhanced feature pyramid, which adeptly amalgamates features extracted across different scales. This strategic enhancement effectively curbs information attrition across various layers, thereby strengthening the feature extraction capabilities of the foundational network. Furthermore, we confront the issue of excessive classification loss in image object detection tasks by introducing the gradient density loss function, designed to mitigate classification discrepancies. Empirical results unequivocally demonstrate the efficacy of our approach in enhancing the detection of multi-scale objects within images. When evaluated across benchmark datasets, including MS COCO 2017, MS COCO 2014, Pascal VOC 2007, and Pascal VOC 2012, our method achieves impressive average precision scores of 39.4%, 42.0%, 51.5%, and 49.9%, respectively. This performance clearly outperforms alternative state-of-the-art methods in the field. This research not only contributes to the evolving landscape of computer vision and object detection but also has practical implications for a wide range of applications, aligning with the transformative trends in the automotive industry and security technologies.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001119950700001  
Indexed Date:  
2023-12-19  
Institution Address:  
Longyan Univ, Sch Phys & Mech & Elect Engn, Longyan 364012, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Sch Software Engn, Xian 710049, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Inst Artificial Intelligence & Robot, Xian 710049, Peoples R China  
Institution:  
Institution:  
Institution:

### 271. Terahertz image enhancement based on a multiscale feature extraction

摘要: The development and application of terahertz (THz) waves hold great potential in military, industrial, and biomedical fields. Terahertz time-domain spectroscopy (THz-TDS) imaging systems capture a sample's time-domain spectral signal to achieve imaging through spectral analysis for intensity and phase information. Challenges in terahertz imaging include spatial diffraction limits, poor image contrast and clarity due to atmospheric water molecule absorption, and Gaussian and impulse noise. This study utilizes a generative adversarial network structure in deep learning models to enhance THz image quality by providing improved denoising and resolution. Through the integration of certain encoder and decoder concepts and introduction of pyramid pooling residual dense block module for feature fusion extraction on low-resolution images, a super-resolution network is designed and employed on selected THz images of deformed metal. Multiple standards are introduced for algorithm performance evaluation. Our experimental results demonstrate that compared with bicubic, super-resolution generative adversarial networks (SRGAN), and residual dense network (RDN) algorithms, our algorithm effectively improves image resolution, and removes noise while preserving high-frequency details without introducing unnecessary high-frequency artifacts. (c) 2024 Optica Publishing Group under the terms of the Optica Open Access Publishing Agreement  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1094-4087  
入藏号:  
WOS:001318757600001  
PubMed ID:  
39572999  
Indexed Date:  
2024-10-05  
Institution Address:  
Chongqing Univ Posts & Telecommun, Sch Optoelect, Chongqing 400065, Peoples R China  
Institution Address:  
Sichuan Univ Arts & Sci, Sch Chengdu Res Inst, Sichuan 635000, Peoples R China  
Institution Address:  
Univ Glasgow, James Watt Sch Engn, Glasgow G12 8QQ, Scotland  
Institution Address:  
Jinan Univ, Inst Nanophoton, Guangzhou 511443, Peoples R China  
Institution Address:  
Univ Elect Sci & Technol China UESTC, Chengdu 610054, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 272. Single-Image Defogging Algorithm Based on Improved Cycle-Consistent

摘要: With the wave of artificial intelligence and deep learning sweeping the world, there are many algorithms based on deep learning for image defog research. However, there is still serious color distortion, contrast reduction, incomplete fog removal, and other problems. To solve these problems, this paper proposes an improved image defogging network based on the traditional cycle-consistent adversarial network. We add the self-attention module and atrous convolution multi-scale feature fusion module on the basis of the traditional CycleGAN network to enhance the feature extraction capability of the network. The perceptual loss function is introduced into the loss function of the model to enhance the texture sense of the generated image. Finally, by comparing several typical defogging algorithms, the superiority of the defogging model proposed in this paper is proved qualitatively and quantitatively. Among them, on the indoor synthetic data set, the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measurement (SSIM) of the network designed by us can reach 23.22 and 0.8809, respectively. On the outdoor synthetic data set, the PSNR and SSIM of our designed network can be as high as 25.72 and 0.8859, respectively. On the real data set, the PSNR and SSIM of our designed network can reach 21.02 and 0.8166, respectively. It is proved that the defogging network in this paper has good practicability and universality.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
被引频次合计:  
6  
eISSN:  
2079-9292  
入藏号:  
WOS:000996953500001  
Indexed Date:  
2023-06-12  
Institution Address:  
Harbin Univ Sci & Technol, Sch Measurement Control Technol & Commun Engn, Harbin 150080, Peoples R China  
Institution:

### 273. Face Recognition Model Optimization Research Based on Embedded Platform

摘要: The development of information technology has promoted the expansion of the application field of facial recognition technology. Its mainstream recognition methods rely on deep learning algorithms for calculation, but the problem of large data computation brought by its system makes it difficult to apply to embedded platform devices. As a result, this study focuses on improving recognition systems built on lightweight backend networks and builds an embedded platform system environment using multi-scale feature fusion, anchor box size optimization, the addition of channel attention mechanism weighted features, affine face alignment, and file compilation. The experimental results showed that when the number of iterations was 300, the loss value (0.46) of the improved embedded algorithm was much smaller than that of other comparison algorithms (1.42, 1.73, 2.01), and its ACC value (0.924) was significantly better than other comparison algorithms (0.915, 0.909, 0.894). The minimum system testing time consumedwas7 ms. This deep learning embedded facial recognition algorithm has high recognition accuracy, a fast running speed, and is less limited by environmental conditions and data types. It is ideally suited for use in embedded hardware devices, broadening the scope of equipment matching and facial recognition algorithms' applications. As a result, it is better suited to satisfy the demands of embedded devices and massive data processing jobs.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001030572500001  
Indexed Date:  
2023-08-02  
Institution Address:  
Shenyang Univ Technol, Sch Informat Sci & Engn, Shenyang 110870, Peoples R China  
Institution Address:  
Neusoft Corp, Shenyang 110179, Peoples R China  
Institution:  
Institution:  
Neusoft Corp

### 274. Deformable YOLOX: Detection and Rust Warning Method of Transmission Line

摘要: Dense distribution and significant size difference of transmission line connecting fittings are difficult to maintain, and long-term exposure to the outdoor environment is vulnerable to adverse environmental effects of rust failure. The common image processing methods and deep learning algorithms are not competent for this kind of dense small-target detection task, so the target detection model based on an image processing hierarchical algorithm is proposed in this article, which uses anchor-free and decoupled head design ideas, through ASFF multiscale information feature fusion strategy and ECA + VariFocal Loss interactive saliency area capture strategy to construct a dense small-target detection network suitable for a complex environment. The experimental results show that the comprehensive performance of Deformable YOLOX is superior to 13 current advanced target detection algorithms. Compared with the baseline model, Deformable YOLOX can better understand the multiscale semantic information of the image and learn the small details that are more difficult to distinguish. Combined with a target detection algorithm, an early warning algorithm for rust grade assessment of connecting fittings is proposed, and an online monitoring system is designed, which has practical engineering application value.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
16  
被引频次合计:  
18  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000930578800003  
Indexed Date:  
2023-03-22  
Institution Address:  
Xian Polytech Univ, Sch Elect & Informat, Xian 710048, Peoples R China  
Institution:

### 275. Synthetic aperture optical image restoration based on multi-scale

摘要: With the wide applications of high-resolution imaging technology in topographic mapping, astronomical observation, and military reconnaissance and other fields, the requirements for imaging resolution of optical system are becoming higher and higher. According to the diffraction limit and Rayleigh criterion, the imaging resolution of the optical system is proportional to the size of the aperture of the system, but affected by the material and the processing of the optical component: the single aperture of the optical system cannot be infinitely enlarged. Therefore the synthetic aperture technology is proposed to replace the single large aperture optical system. Owing to the effect of sub-aperture arrangement and light scattering, the imaging of synthetic aperture optical system will be degraded because of insufficient light area and phase distortion. The traditional imaging restoration algorithm of synthetic aperture optical system is sensitive to noise, overly relies on degraded model, requires a lot of manually designed models, and has poor adaptability. To solve this problem, a multiscale feature enhancement method of restoring the synthetic aperture optical image is proposed in this work. UNet is used to obtain multi-scale feature, and self-attention in mixed domain is used to improve the ability of of the network to extract the features in space and channel. Multi-scale feature fusion module and feature enhancement module are constructed to fuse the information between features on different scales. The information interaction mode of the codec layer is optimized, the attention of the whole network to the real structure of the original image is enhanced, and the artifact interference caused by ringing is avoided in the process of restoration. The final experimental results are 1.51%, 4.42% and 5.22% higher than those from the advanced deep learning algorithms in the evaluation indexes of peak signal-to-noise ratio, structural similarity and perceived similarity, respectively. In addition, the method presented in this work has a good restoration effect on the degraded images to different degrees of synthetic aperture, and can effectively restore the degraded images and the images with abnormal light, so as to solve the problem of imaging degradation of synthetic aperture optical system. The feasibility of deep learning method in synthetic aperture optical image restoration is proved.  
摘要:  
受物理孔径大小和光线散射等影响,合成孔径光学系统成像因通光面积不足和相位失真而出现降质模糊.传统合成孔径光学系统成像复原算法对噪声敏感,过于依赖退化模型,自适应性差.对此提出一种基于生成对抗网络的光学图像复原方法,采用U-Net结构获取图像多级尺度特征,利用基于自注意力的混合域注意力提高网络在空间、通道上的特征提取能力,构造多尺度特征融合模块和特征增强模块,融合不同尺度特征间的信息,优化了编解码层的信息交互方式,增强了整体网络对原始图像真实结构的关注力,避免在复原过程中被振铃现象产生的伪影干扰.实验结果表明,与其他现有方法相比,该方法在峰值信噪比、结构相似性和感知相似度评估指标上分别提高了1.51%, 4.42%和5.22%,有效解决合成孔径光学系统成像结果模糊退化的问题.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1000-3290  
入藏号:  
CSCD:7693248  
Indexed Date:  
2024-06-21  
Institution Address:  
无锡学院;;南京信息工程大学电子信息工程学院, 江苏省集成电路可靠性技术及检测系统工程研究中心;;, 无锡;;南京, ;; 214105;;210044  
Institution Address:  
南京信息工程大学电子信息工程学院, 南京, 江苏 210044, 中国  
Institution Address:  
Wuxi University;;School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Jiangsu Province Engineering Research Center of Integrated Circuit Reliability Technology and Testing System;;, Wuxi;;Nanjing, ;; 214105;;210044  
Institution Address:  
School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China  
Institution:  
无锡学院;;南京信息工程大学电子信息工程学院  
Institution:  
南京信息工程大学电子信息工程学院  
Institution:  
Wuxi University;;School of Electronic and Information Engineering, Nanjing University of Information Science and Technology  
Institution:  
School of Electronic and Information Engineering, Nanjing University of Information Science and Technology

### 276. GSBF-YOLO: a lightweight model for tomato ripeness detection in natural

摘要: Accurate tomato ripeness detection is essential for optimizing harvest timing and maximizing yield. Deep learning-based object detection has proven effective in this task. However, many existing algorithms have numerous parameters and substantial computational demands, making them unsuitable for agricultural environments with limited computational resources. Additionally, accurate detection becomes challenging with overlapping fruits, leaf occlusion, or complex backgrounds. To address these issues, this paper proposes a lightweight detection model, GSBF-YOLO. This model designs the GSim module to reduce parameters while maintaining detection accuracy. The C3Ghost module further reduces parameter count by replacing the traditional C3 module. The PANet multi-scale feature fusion network in the neck is replaced with the Bi-directional Feature Pyramid Network (BiFPN), which adjusts weights based on the importance of input features. Lastly, the fine-tuned FocalEIOU Loss function is used to calculate the bounding box regression loss, enhancing the model's ability to adjust the weights of high-quality anchor boxes for better detection of targets in occlusion scenarios. Experimental results show that GSBF-YOLO reduces parameters and computational load by 42% and 45%, respectively, while mean Average Precision (mAP) increases by 1.9% and 1.6% on two datasets. The model achieves 110 Frames Per Second (FPS), meeting real-time detection requirements, and has fewer parameters and higher accuracy compared to models like YOLOv8. The research indicates that the proposed lightweight model can effectively detect tomato ripeness in natural environments.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001400424400001  
Indexed Date:  
2025-01-25  
Institution Address:  
Qilu Univ Technol, Key Lab Comp Power Network & Informat Secur, Minist Educ,Shandong Comp Sci Ctr, Natl Supercomp Ctr Jinan,Shandong Acad Sci, Jinan 250014, Shandong, Peoples R China  
Institution Address:  
City Univ Macau, Fac Data Sci, Macau 999078, Peoples R China  
Institution Address:  
Shandong Fundamental Res Ctr Comp Sci, Shandong Prov Key Lab Ind Network & Informat Syst, Jinan 250014, Peoples R China  
Institution Address:  
Qingdao Agr Univ, Commun Coll, Qingdao 266109, Peoples R China  
Institution:  
Institution:  
Institution:  
Shandong Fundamental Res Ctr Comp Sci  
Institution:

### 277. Ai-aided diagnosis of oral X-ray images of periapical films based on

摘要: Oral X-ray images provide a useful technical means by which dentists examine teeth for dental problems, but the diagnostic process is defective due to its over-reliance on dentists' subjective judgments, lack of objective criteria, etc. In this context, this study examined the AI-aided diagnosis of periapical films based on deep learning..Based on YOLOv7-X, a YOLO-DENTAL network architecture was used to detect dental caries, dental defects, periapical lesions, and coronal restorations in periapical films. Firstly, the coordinate attention (CA) mechanism was introduced into the backbone feature extraction network, and a backbone-CA structure was presented to enhance the feature extraction capability of the network. Secondly, a simplified Bi-FPN structure was put forward and applied to the feature fusion part of the network to effectively improve its multi-scale feature fusion effect. Thirdly, the existing anchor-based detection head was replaced by an anchor-free decoupled head to simplify operational parameters while improving the generalized detection capability of the model over lesion regions. In the loss function part, existing CIoU loss was replaced by SIoU loss, a border loss function containing direction information. The focal loss containing a weight factor was introduced in calculating confidence loss as a substitute for the existing binary cross entropy loss function to balance positive and negative samples. Meanwhile, a study of ablation experiment was completed. The results validated the positive gain effect of each optimization strategy on the model. The final YOLO-DENTAL network structure exhibited an mAP value of 86.81%, higher than that of YOLOv7-X (79.95%). The effect of aided diagnosis was well achieved.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0141-9382  
eISSN:  
1872-7387  
入藏号:  
WOS:001172106800001  
Indexed Date:  
2024-03-13  
Institution Address:  
Beijing Jiaotong Univ, Natl Phys Expt Teaching Demonstrat Ctr, Sch Phys Sci & Engn, Key Lab Educ,Minist Luminescence & Opt Informat Te, Beijing 100044, Peoples R China  
Institution Address:  
Peking Univ, Sch & Hosp Stomatol, Dept Gen Dent 2, Beijing 100081, Peoples R China  
Institution Address:  
Natl Ctr Stomatol, Beijing 100081, Peoples R China  
Institution Address:  
Natl Clin Res Ctr Oral Dis, Beijing 100081, Peoples R China  
Institution Address:  
Natl Engn Lab Digital & Mat Technol Stomatol, Beijing 100081, Peoples R China  
Institution Address:  
Beijing Key Lab Digital Stomatol, Beijing 100081, Peoples R China  
Institution Address:  
Minist Hlth, Res Ctr Engn & Technol Computerized Dent, Beijing 100081, Peoples R China  
Institution Address:  
NMPA Key Lab Dent Mat, Beijing 100081, Peoples R China  
Institution Address:  
Beijing Chaoyang Tongwen Forenign Language Sch, Sch Automat, Beijing 100081, Peoples R China  
Institution Address:  
Beijing Jiaotong Univ, Sch Comp & Informat Technol, Beijing 100044, Peoples R China  
Institution:  
Institution:  
Institution:  
Natl Ctr Stomatol  
Institution:  
Natl Clin Res Ctr Oral Dis  
Institution:  
Natl Engn Lab Digital & Mat Technol Stomatol  
Institution:  
Beijing Key Lab Digital Stomatol  
Institution:  
Minist Hlth  
Institution:  
NMPA Key Lab Dent Mat  
Institution:  
Beijing Chaoyang Tongwen Forenign Language Sch  
Institution:

### 278. MARA-YOLO: An Efficient Method for Multiclass Personal Protective

摘要: Personal Protective Equipment (PPE) plays a crucial role in protecting workers from unpredictable physical threats and reducing fatal occupational injury rates. However, PPE is not always worn by workers in all cases, which poses a challenge to site management. In this context, object detection methods based on deep learning have been used to strengthen site management. To address the challenges posed by limited PPE detection performance under complex environmental conditions, we develop the MARA-YOLO object detection model, which is designed to balance speed and accuracy. Firstly, based on YOLOv8-s, we introduce a modified re-parameterizable backbone, which consists of MobileOne-S0 and an efficient downsampling block known as the Attentional Space-to-Depth Block (AS-Block). Subsequently, we propose the R-C2F module, which fused feature maps from diverse receptive fields and enhances the model's sensitivity to the texture information of objects, as well as its ability to capture information from varying depths. Build upon R-C2F and Adaptively Spatial Feature Fusion (ASFF), a multi-scale feature fusion module RASFF is further introduced to mitigate inconsistent multi-scale outputs in the model. Finally, a dedicated dataset consisting of 2750 images covering 9 categories is constructed. The ablation experiments demonstrate that compared to the baseline, MARA-YOLO achieves a 6.7% improvement in AP(50) and a 10.2% improvement in AP(75) on the proposed KSE-PPE dataset. In the comparative experiments, MARA-YOLO achieves a mean average precision (mAP) of 74.7% on the KSE-PPE dataset, surpassing other lightweight state-of-the-art models by more than 4.95%.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
2169-3536  
入藏号:  
WOS:001173071400001  
Indexed Date:  
2024-03-14  
Institution Address:  
Kunming Univ Sci & Technol, Fac Mech & Elect Engn, Kunming 650031, Yunnan, Peoples R China  
Institution Address:  
Nanchang Univ, Sch Econ & Management, Nanchang 330036, Jiangxi, Peoples R China  
Institution:  
Institution:

### 279. A Novel Deep-Learning-Based QoS Prediction Model for Service

摘要: With the rapid development of service computing, the demand for service recommendation is increasing. Quality of Service (QoS) prediction has been one of the key challenges for service recommendation. Existing deep learning-based methods have been proposed for QoS prediction, but further improvement of their neural network structures is still needed to improve the prediction accuracy. This work introduces multi-stage multi-scale feature fusion with individual evaluations to a deep learning model for accurate QoS prediction. In our model, non-negative matrix factorization is used to extract three-scale (i.e., global, local, and individual) features; distance similarity is exploited to find similar users and services; a multi-stage deep neural network is designed to fuse multi-scale features, where individual evaluations are input to each stage to correct QoS features. Finally, our model is compared with often-cited prediction methods, and the experimental results show that it can more accurately predict QoS than its peers.Note to Practitioners-Accurate QoS prediction is very helpful to recommend the most suitable services to users among many similar services. Affected by the sparsity of historical data, the accuracy of existing QoS prediction methods is often limited. The multi-scale features of users and services can be used to improve prediction accuracy. This work proposes a new QoS prediction method to do so. Specifically, it first extracts global and individual features through non-negative matrix factorization and uses distance similarity to obtain local features. Then, it proposes a new deep neural network that fuses the extracted multi-scale features in each learning stage, thereby improving QoS prediction for services recommendation.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
14  
ISSN:  
1545-5955  
eISSN:  
1558-3783  
入藏号:  
WOS:001006381500001  
Indexed Date:  
2023-06-27  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Minist Educ, Engn Res Ctr Digital Forens, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Comp Sci, Nanjing 210044, Peoples R China  
Institution Address:  
New Jersey Inst Technol, Helen & John C Hartmann Dept Elect & Comp Engn, Newark, NJ 07102 USA  
Institution Address:  
King Abdulaziz Univ, KA CARE Energy Res & Innovat Ctr, Jeddah 21481, Saudi Arabia  
Institution:  
Institution:  
Institution:  
Institution:

### 280. Predicting length of stay in ICU and mortality with temporal dilated

摘要: In healthcare, Intensive Care Unit (ICU) bed management is a necessary task because of the limited budget and resources. Predicting the remaining Length of Stay (LoS) in ICU and mortality can assist clinicians in managing ICU beds efficiently. This study proposes a deep learning method based on several successive Temporal Dilated Separable Convolution with Context-Aware Feature Fusion (TDSC-CAFF) modules, and a multi-view and multi -scale feature fusion for predicting the remaining LoS and mortality risk for ICU patients. In each TDSC-CAFF module, temporal dilated separable convolution is used to encode each feature separately, and context-aware feature fusion is proposed to capture comprehensive and context-aware feature representations from the input time-series features, static demographics, and the output of the last TDSC-CAFF module. The CAFF outputs of each module are accumulated to achieve multi-scale representations with different receptive fields. The outputs of TDSC and CAFF are concatenated with skip connection from the output of the last module and the original time-series input. The concatenated features are processed by the proposed Point-Wise convolution-based Attention (PWAtt) that captures the inter-feature context to generate the final temporal features. Finally, the final temporal features, the accumulated multi-scale features, the encoded diagnosis, and static demographic features are fused and then processed by fully connected layers to obtain prediction results. We evaluate our proposed method on two publicly available datasets: eICU and MIMIC-IV v1.0 for LoS and mortality prediction tasks. Experimental results demonstrate that our proposed method achieves a mean squared log error of 0.07 and 0.08 for LoS prediction, and an Area Under the Receiver Operating Characteristic Curve of 0.909 and 0.926 for mortality prediction, on eICU and MIMIC-IV v1.0 datasets, respectively, which outperforms several state-of-the-art methods.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
4  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
7  
被引频次合计:  
7  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:000900142700004  
PubMed ID:  
36371901  
Indexed Date:  
2023-01-15  
Institution Address:  
Cent South Univ, Sch Comp Sci & Engn, Hunan Prov Key Lab Bioinformat, Changsha 410083, Hunan, Peoples R China  
Institution Address:  
Sanaa Univ, Fac Comp & Informat Technol, Sanaa, Yemen  
Institution:  
Institution:  
Sanaa Univ

### 281. TG-Net: A Physically Interpretable Deep Learning Forecasting Model for

摘要: Thunderstorm gusts are a common and hazardous type of severe convective weather, characterized by a small spatial scale, short duration, and significant destructive power. They often lead to severe disasters, highlighting the critical importance of their accurate forecasting. Previous studies have explored the environmental factors and spatiotemporal distribution characteristics of thunderstorm gusts, highlighting the need for improved forecasting methods. In recent years, artificial intelligence techniques have shown promise in enhancing the accuracy of thunderstorm gust forecasting, with various machine learning algorithms and models having been developed. This paper proposes a multiscale feature fusion module called Thunderstorm Gusts Block (TG-Block) and a deep learning model named Thunderstorm Gusts net (TG-net) based on the Attention U-net and TG-TransUnet models, and employs interpretable methods such as Integrated Gradient, Deep Learning Importance Features, and Shapley Additive exPlanations to validate the model's practical relevance and reliability. The analysis of feature importance underscores the model's ability to capture key thermodynamic and multiscale weather characteristic information for thunderstorm gust nowcasting. It is, however, worth emphasizing that these conclusions are only based on a limited number of thunderstorm gust examples, and the evaluation results may be affected by specific weather types and sample sizes. Nonetheless, TG-net has been put into real-time operation at the Institute of Urban Meteorology, and we will continue to rigorously validate its performance and make any necessary optimizations and enhancements based on feedback to ensure the robustness and stability of the model.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2095-6037  
eISSN:  
2198-0934  
入藏号:  
WOS:001439661700010  
Indexed Date:  
2025-03-15  
Institution Address:  
Ocean Univ China, Fac Informat Sci & Engn, Qingdao 266100, Peoples R China  
Institution Address:  
China Meteorol Adm, Inst Urban Meteorol, Beijing 100089, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Collaborat Innovat Ctr Forecast & Evaluat Meteorol, Nanjing 210044, Peoples R China  
Institution:  
Institution:  
Institution:

### 282. Seismic waveform classification based on a multiscale attention residual

摘要: A total of 117 seismic events (54 natural earthquakes and 63 explosions) that occurred in the Capital Circle Region (Shunyi, Beijing,Sanhe,and Hebei) from February 2010 to December 2016 were selected in this paper. The multiscale attention residual network was proposed and used to classify the waveforms of earthquakes and explosions. The original seismic waveform was simply preprocessed and intercepted into seismic time series data with the same length,which was directly used as the input of the network model. Then, the deep neural network with the residual module was selected as the basic network. The step of advanced extraction of time-domain waveform features as the input of classification algorithm in traditional waveform classification can be omitted by using the automatic feature extraction ability of the deep neural network. Next, the efficient channel attention mechanism was integrated and improved, after which information from the spatial dimension was integrated into the channel information, thus optimizing the network's attention to key information and resulting in better concentration on essential features. Finally, the multiscale feature fusion was performed using spatial pyramid pooling instead of maximum pooling to obtain more feature information. Ultimately,a multiscale attention residual network was formed. Experimental results show that the highest classification accuracy of the multiscale attention residual network is 97.11%, and the average classification accuracy is 96.53%. The results demonstrate the effectiveness of this approach in seismic waveform classification and provide a new optional approach for seismic-source type identification.  
摘要:  
选用2010年2月-2016年12月发生在北京顺义及河北三河等首都圈邻近区域的117个地震事件(包括54个天然地震事件和63个非天然地震事件--爆炸事件)作为研究对象,利用文章所提出的多尺度注意残差网络对其中的天然地震事件和爆炸事件波形进行二分类。首先,对原始地震波形进行简单预处理并截取成相同长度的地震时序数据,直接将其作为网络模型的输入;其次,选用含有残差模块的深度神经网络作为基础网络,利用深度神经网络对特征的自动提取能力,省略了传统波形分类需要提前提取时域波形的特征作为分类算法输入的步骤;然后,融合通道注意力机制(ECA)并对其进行改进,将空间维度的信息融入通道信息,优化了网络对关键信息的关注,更好地聚焦重要特征;最后,使用空间金字塔池化代替最大池化进行多尺度特征融合,得到更多的特征信息,构成多尺度注意残差网络。实验结果表明,最高分类准确率为97.11%,平均分类准确率为96.53%,证明了多尺度注意残差网络在地震波形分类任务中的有效性,为震源类型识别工作提供了一种新的方法。  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1000-0844  
入藏号:  
CSCD:7757705  
Indexed Date:  
2024-09-06  
Institution Address:  
广西师范大学计算机科学与工程学院/软件学院, 桂林, 广西 541004, 中国  
Institution Address:  
College of Computer Science and Engineering & College of Software,Guangxi Normal University, Guilin, Guangxi 541004, China  
Institution:  
广西师范大学计算机科学与工程学院/软件学院  
Institution:  
College of Computer Science and Engineering & College of Software,Guangxi Normal University

### 283. Light weight object detector based on composite attention residual

摘要: The object detector based on deep learning has received extensive attention, but the high computational cost has become an obstacle to its large-scale application. It is a great challenge for object detection to further reduce the hardware requirements on the premise of ensuring high detection accuracy. We propose a one-stage lightweight object detector and a new regression loss. In this method, ResNet is improved and combined with attention mechanism to ensure the maximum integrity of feature information with fewer parameters; The multi-scale feature fusion network is improved to reduce the reasoning complexity of the structure. In addition, the bounding box regression loss is improved, and the specific position of the bounding box is adjusted by considering the balance of multiple factors in the regression process. The experimental results show that: 1) the combination of most detectors and improved loss can further improve the performance of detectors; 2) As a whole, our improved network and loss can give consideration to both speed and accuracy on Pascal VOC and COCO; 3) in the addition of other new training tricks such as DropBlock and Mosaic, we can achieve better overall performance on the coco test development set, 38.42 AP (average accuracy) at 40.3 FPS. (C) 2022 Elsevier B.V. All rights reserved.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000830184400012  
Indexed Date:  
2022-08-06  
Institution Address:  
Tianjin Univ Technol, Tianjin Key Lab Control Theory & Applicat Complic, Tianjin 300384, Peoples R China  
Institution Address:  
China North Ind Grp 210 Res Inst, Beijing 100089, Peoples R China  
Institution Address:  
Univ South Africa, Dept Elect Engn, ZA-1710 Florida, South Africa  
Institution:  
Institution:  
China North Ind Grp 210 Res Inst  
Institution:

### 284. Enhancement-suppression driven lightweight fine-grained micro-expression

摘要: Micro-expressions are short-lived and authentic emotional expressions used in several fields such as deception detection, criminal analysis, and medical diagnosis. Although deep learning-based approaches have achieved outstanding performance in micro-expression recognition, the recognition performance of lightweight networks for terminal applications is still unsatisfactory. This is mainly because existing models either excessively focus on a single region or lack comprehensiveness in identifying various regions, resulting in insufficient extraction of fine-grained features. To address this problem, this paper proposes a lightweight micro-expression recognition framework -Lightweight Fine-Grained Network (LFGNet). The proposed network adopts EdgeNeXt as the backbone network to effectively combine local and global features, as a result, it greatly reduces the complexity of the model while capturing micro-expression actions. To further enhance the feature extraction ability of the model, the Enhancement-Suppression Module (ESM) is developed where the Feature Suppression Module(FSM) is used to force the model to extract other potential features at deeper layers. Finally, a multi- scale Feature Fusion Module (FFM) is proposed to weigh the fusion of the learned features at different granularity scales for improving the robustness of the model. Experimental results, obtained from four datasets, demonstrate that the proposed method outperforms already existing methods in terms of recognition accuracy and model complexity.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1047-3203  
eISSN:  
1095-9076  
入藏号:  
WOS:001401677700001  
Indexed Date:  
2025-01-28  
Institution Address:  
Shandong Technol & Business Univ, Sch Informat & Elect Engn, Yantai 264000, Shandong, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automat, Beijing 100000, Peoples R China  
Institution:  
Institution:

### 285. Seismic P-Wave First-Arrival Picking Model Based on Spatiotemporal

摘要: Aiming at the problems of low accuracy and poor robustness of the existing earthquake first-arrival picking algorithm,a seismic P-wave arrival picking network based on deep learning is designed.This network is encoder-decoder structure,which can identify seismic signal sequence point by point.The encoder uses multi-scale feature extractor for feature extraction and fusion of input data to improve feature utilization ratio.The multi-scale residual structure is used to deeply mine the hidden feature information in the data to improve the nonlinear fitting ability of the model.Then,the spatiotemporal attention mechanism is added to the decoder to improve the network's perception of the first-arrival features.Finally,a deep coding feature fusion module is proposed to effectively avoid the pollution of feature sequence while ensuring the integrity of features.The experimental results show that under the three error thresholds of 0.1 s,0.2 s and 0.3 s,the picking hit rate of the proposed network are 75.04%,94.6% and 97.37%,respectively,the mean absolute error and mean square error are 0.092 s and 0.036.Compared with the existing traditional and deep learning first-arrival picking methods,it has higher P-wave first-arrival picking accuracy.  
摘要:  
针对现有地震到时拾取算法精度较低、鲁棒性较差等问题,设计了一种基于深度学习的地震P波到时拾取网络,该网络为编解码结构,可实现地震波形序列的逐点预测。网络编码器对输入数据进行多尺度特征提取与融合,提高特征利用率;利用多尺度残差结构深度挖掘数据中隐藏特征信息,提升模型非线性拟合能力;在解码网络中加入时空注意力机制,提高网络对到时特征的感知能力;提出深层编码特征融合模块,在保证特征完整性的同时有效避免融合特征过程中出现的特征序列污染问题。实验结果表明,提出的网络在0.1 s、0.2 s、0.3 s三个误差阈值下,拾取命中率分别为75.04%、94.6%、97.37%,平均绝对误差和均方误差为0.092 s、0.036,相比现有传统方法与深度学习到时拾取方法,具有更高的P波到时拾取精度。  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1002-8331  
入藏号:  
CSCD:7435784  
Indexed Date:  
2023-07-07  
Institution Address:  
太原理工大学大数据学院, 太原, 山西 030000, 中国  
Institution Address:  
太原理工大学软件学院, 太原, 山西 030000, 中国  
Institution Address:  
太原理工大学信息与计算机学院, 太原, 山西 030000, 中国  
Institution Address:  
College of Data Science,Taiyuan University of Technology, Taiyuan, Shanxi 030000, China  
Institution Address:  
College of Software,Taiyuan University of Technology, Taiyuan, Shanxi 030000, China  
Institution Address:  
College of Information and Computer,Taiyuan University of Technology, Taiyuan, Shanxi 030000, China  
Institution:  
太原理工大学大数据学院  
Institution:  
太原理工大学软件学院  
Institution:  
太原理工大学信息与计算机学院  
Institution:  
College of Data Science,Taiyuan University of Technology  
Institution:  
College of Software,Taiyuan University of Technology  
Institution:  
College of Information and Computer,Taiyuan University of Technology

### 286. Flood Detection in Dual-Polarization SAR Images Based on Multi-Scale

摘要: The proliferation of massive polarimetric Synthetic Aperture Radar (SAR) data helps promote the development of SAR image interpretation. Due to the advantages of powerful feature extraction capability and strong adaptability for different tasks, deep learning has been adopted in the work of SAR image interpretation and has achieved good results. However, most deep learning methods only employ single-polarization SAR images and ignore the water features embedded in multi-polarization SAR images. To fully exploit the dual-polarization SAR data and multi-scale features of SAR images, an effective flood detection method for SAR images is proposed in this paper. In the proposed flood detection method, a powerful Multi-Scale Deeplab (MS-Deeplab) model is constructed based on the dual-channel MobileNetV2 backbone and the classic DeeplabV3+ architecture to improve the ability of water feature extraction in SAR images. Firstly, the dual-channel feature extraction backbone based on the lightweight MobileNetV2 separately trains the dual-polarization SAR images, and the obtained training parameters are merged with the linear weighting to fuse dual-polarization water features. Given the multi-scale space information in SAR images, then, a multi-scale feature fusion module is introduced to effectively utilize multi-layer features and contextual information, which enhances the representation of water features. Finally, a joint loss function is constructed based on cross-entropy and a dice coefficient to deal with the imbalanced categorical distribution in the training dataset. The experimental results on the time series of Sentinel-1A SAR images show that the proposed method for flood detection has a strong ability to locate water boundaries and tiny water bodies in complex scenes. In terms of quantitative assessment, MS-Deeplab can achieve a better performance compared with other mainstream semantic segmentation models, including PSPNet, Unet and the original DeeplabV3+ model, with a 3.27% intersection over union (IoU) and 1.69% pixel accuracy (PA) improvement than the original DeeplabV3+ model.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
15  
eISSN:  
2072-4292  
入藏号:  
WOS:000876118700001  
Indexed Date:  
2022-11-11  
Institution Address:  
Hangzhou Dianzi Univ, Sch Commun Engn, Hangzhou 310018, Peoples R China  
Institution Address:  
Natl Univ Def Technol, Natl Key Lab Sci & Technol Automat Target Recogni, Changsha 410073, Peoples R China  
Institution:  
Institution:

### 287. WaveFusionNet: Infrared and visible image fusion based on multi-scale

摘要: To merge complementary information from multimodal images, such as thermal saliency from infrared images and texture details from visible images, traditional multi-scale transform-based methods have been extensively studied, with deep learning-based methods gaining significant popularity in recent years. However, there has been limited research on optimally combining the advantages of these two categories in fusion. In this paper, we propose a novel infrared and visible image fusion (IVIF) framework, WaveFusionNet, which integrates precise frequency feature decomposition from the discrete wavelet transform (DWT) with the comprehensive feature extraction from the multi-scale encoder. Firstly, we train an encoder-decoder network for multi- scale feature extraction and image reconstruction. DWT is used for down-sampling with minimal information loss by decomposing extracted features into low and high-frequency sub-bands. Next, a dual-band feature fusion (DBFF) module is trained to merge these sub-bands by integrating a spatial feature transform-based sub-network for low-frequency fusion and a maximum absolute value selection strategy for fusing high- frequencies. Finally, all fused sub-bands are fed into the pre-trained decoder to reconstruct the final image. Experimental results on three benchmark datasets (TNO, Roadscene, and MSRS) demonstrate that the proposed fusion method outperforms recent IVIF methods in both quantitative assessment and visual perception while maintaining competitive time complexity.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0030-4018  
eISSN:  
1873-0310  
入藏号:  
WOS:001316981500001  
Indexed Date:  
2024-09-27  
Institution Address:  
Tianjin Univ, Sch Microelect, Tianjin 300072, Peoples R China  
Institution Address:  
Univ British Columbia, Dept Comp Sci Math Phys & Stat, Okanagan Campus, Kelowna, BC V1V 1V7, Canada  
Institution:  
Institution:

### 288. Brain metastasis magnetic resonance imaging-based deep learning for

摘要: Background: The preoperative identification of epidermal growth factor receptor ( EGFR ) mutations and subtypes based on magnetic resonance imaging (MRI) of brain metastases (BM) is necessary facilitate individualized therapy. This study aimed to develop a deep learning model to preoperatively detect EGFR mutations and identify the location of EGFR mutations in patients with non-small cell lung cancer (NSCLC) and BM. Methods: We included 160 and 72 patients who underwent contrast-enhanced T1-weighted (T1w-CE) and T2-weighted (T2W) MRI at Liaoning Cancer Hospital and Institute (center 1) and Shengjing Hospital China Medical University (center 2) to form a training cohort and an external validation cohort, respectively. A multiscale feature fusion network (MSF-Net) was developed by adaptively integrating features based different stages of residual network (ResNet) 50 and by introducing channel and spatial attention modules. The external validation set from center 2 was used to assess the performance of MSF-Net and to compare it with that of handcrafted radiomics features. Receiver operating characteristic (ROC) curves, accuracy, precision, recall, and F1-score were used to evaluate the effectiveness of the models. Gradient-weighted class activation mapping (Grad-CAM) was used to demonstrate the attention of the MSF-Net model. Results: The developed MSF-Net generated a better diagnostic performance than did the handcrafted radiomics in terms of the microaveraged area under the curve (AUC) (MSF-Net: 0.91; radiomics: 0.80) and macroaveraged AUC (MSF-Net: 0.90; radiomics: 0.81) for predicting EGFR mutations and subtypes. Conclusions: This study provides an end-to-end and noninvasive imaging tool for the preoperative prediction of EGFR mutation status and subtypes based on BM, which may be helpful for facilitating individualized clinical treatment plans.  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
2223-4292  
eISSN:  
2223-4306  
入藏号:  
WOS:001292836900014  
PubMed ID:  
39022238  
Indexed Date:  
2024-08-23  
Institution Address:  
China Med Univ, Sch Intelligent Med, 77 Puhe Rd, Shenyang 110122, Peoples R China  
Institution Address:  
Fudan Univ, Sch Informat Sci & Technol, Dept Biomed Engn, Shanghai, Peoples R China  
Institution Address:  
China Med Univ, Liaoning Canc Hosp & Inst, Canc Hosp, Dept Pathol, Shenyang, Peoples R China  
Institution Address:  
China Med Univ, Liaoning Canc Hosp & Inst, Canc Hosp, Dept Radiol, 44 Xiaoheyan Rd, Shenyang 110042, Peoples R China  
Institution Address:  
China Med Univ, Shengjing Hosp, Dept Radiol, 36 Sanhao St, Shenyang 110004, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 289. LMFE-RDD: a road damage detector with a lightweight multi-feature

摘要: Road damage detection using computer vision and deep learning to automatically identify all kinds of road damage is an efficient application in object detection, which can significantly improve the efficiency of road maintenance planning and repair work and ensure road safety. However, due to the complexity of target recognition, the existing road damage detection models usually carry a large number of parameters and a large amount of computation, resulting in a slow inference speed, which limits the actual deployment of the model on the equipment with limited computing resources to a certain extent. In this study, we propose a road damage detector named LMFE-RDD for balancing speed and accuracy, which constructs a Lightweight Multi-Feature Extraction Network (LMFE-Net) as the backbone network and an Efficient Semantic Fusion Network (ESF-Net) for multi-scale feature fusion. First, as the backbone feature extraction network, LMFE-Net inputs road damage images to obtain three different scale feature maps. Second, ESF-Net fuses these three feature graphs and outputs three fusion features. Finally, the detection head is sent for target identification and positioning, and the final result is obtained. In addition, we use WDB loss, a multi-task loss function with a non-monotonic dynamic focusing mechanism, to pay more attention to bounding box regression losses. The experimental results show that the proposed LMFE-RDD model has competitive accuracy while ensuring speed. In the Multi-Perspective Road Damage Dataset, combining the data from all perspectives, LMFE-RDD achieves the detection speed of 51.0 FPS and 64.2% mAP@0.5, but the parameters are only 13.5 M.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0942-4962  
eISSN:  
1432-1882  
入藏号:  
WOS:001260781800002  
Indexed Date:  
2024-07-08  
Institution Address:  
Minnan Normal Univ, Sch Math & Stat, Zhangzhou 363000, Peoples R China  
Institution Address:  
Minnan Normal Univ, Fujian Key Lab Granular Comp & Applicat, Zhangzhou 363000, Peoples R China  
Institution:  
Institution:

### 290. High-resolution building extraction based on the edge-aware network

摘要: Spatial information such as building location and distribution plays an important role in urban dynamic monitoring and urban planning applications. In recent years, deep learning methods have developed rapidly and achieved state-of-the-art performance in building extraction from remote sensing images in a variety of scenarios. However, existing semantic segmentation models pay more attention to global semantic information, emphasize multi-scale feature fusion or set lighter acceptance domains to obtain more global features, and ignore low-level detail features such as edges. Therefore, a new end-to-end deep learning network CEEAU\_Net based on encoder-decoder architecture is designed to add edge sensing module and edge feature extraction module to obtain edge feature information of buildings. The Luxian county area of Luzhou City, Sichuan province is selected for building dataset production, which is located in the Longmenshan seismic zone, with many earthquakes of magnitude three or above, and the scene is complex, so a more accurate building extraction method is needed. Comparison experiments are also conducted with several advanced models on two public datasets, WHU building dataset (WHU) and Massachusetts. Selection of multiple indicators for indicator evaluation of results. CEEAU\_Net achieves the best results in the metrics of overall accuracy, F1-score, Intersection over Union (IoU) and Mean Intersection over Union (MIoU), which suggests that the method proposed in this paper can effectively improve the accuracy of building extraction.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001152477900001  
Indexed Date:  
2024-02-07  
Institution Address:  
Chengdu Univ Technol, Coll Geophys, Chengdu 610059, Peoples R China  
Institution Address:  
Yango Univ, Fujian Key Lab Spatial Informat Percept & Intellig, Fuzhou 350015, Peoples R China  
Institution Address:  
Chengdu Univ Technol, State Key Lab Geohazard Prevent & Geoenvironm Prot, Chengdu 610059, Peoples R China  
Institution:  
Institution:  
Yango Univ  
Institution:

### 291. A novel small object detection algorithm for UAVs based on YOLOv5

摘要: Due to the advances in deep learning, artificial intelligence is widely utilized in numerous areas. Technologies frontier, including computer vision, represented by object detection, have endowed unmanned aerial vehicles (UAVs) with autonomous perception, analysis, and decision-making capabilities. UAVs extensively used in numerous fields including photography, industry and agriculture, surveillance, disaster relief, and play an important role in real life. However, current object detection algorithms encountered challenges when it came to detecting small objects in images captured by UAVs. The small size of the objects, with high density, low resolution, and few features make it difficult for the algorithms to achieve high detection accuracy and are prone to miss and false detections especially when detecting small objects. For the case of enhancing the performance of UAV detection on small objects, a novel small object detection algorithm for UAVs adaptation based on YOLOv5s (UA-YOLOv5s) was proposed. (1) To achieve effective small-sized objects detection, a more accurate small object detection (MASOD) structure was adopted. (2) To boost the detection accuracy and generalization ability of the model, a multi-scale feature fusion (MSF) approach was proposed, which fused the feature information of the shallow layers of the backbone and the neck. (3) To enhance the model stability properties and feature extraction capability, a more efficient and stable convolution residual Squeeze-and-Excitation (CRS)module was introduced. Compared with the YOLOv5s, mAP@0.5 was achieved an impressive improvement of 7.2%. Compared with the YOLOv5l, mAP@0.5 increased by 1.0%, and GFLOPs decreased by 69.1%. Compared to the YOLOv3, mAP@0.5 decreased by 0.2% and GFLOPs by 78.5%. The study's findings demonstrated that the proposed UA-YOLOv5s significantly enhanced the object detection performance of UAVs campared to the traditional algorithms.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
7  
ISSN:  
0031-8949  
eISSN:  
1402-4896  
入藏号:  
WOS:001154784200001  
Indexed Date:  
2024-02-09  
Institution Address:  
Qilu Univ Technol, Shandong Acad Sci, Sch Informat & Automat Engn, Jinan 250353, Shandong, Peoples R China  
Institution Address:  
Dalhousie Univ, Sch Elect & Comp Engn, Halifax, NS B3J 1Z1, Canada  
Institution Address:  
Shandong Runyi Intelligent Technol Co Ltd, Jinan 250002, Peoples R China  
Institution:  
Institution:  
Institution:  
Shandong Runyi Intelligent Technol Co Ltd

### 292. CORNet: Context-Based Ordinal Regression Network for Monocular Depth

摘要: Monocular depth estimation, as one of the fundamental tasks of computer vision, plays a crucial role in three-dimensional (3D) scene understanding and perception. Usually, deep learning methods recover monocular depth maps using continuous regression manners by minimizing the errors between the ground-truth depth and the predicted depth. However, fine depth features may not be fully captured through layer-by-layer coding, which is prone to low spatial resolution depth maps and insufficient details. Furthermore, it usually converges slowly and suffers from unsatisfactory results. To tackle these issues, we propose a novel model, named context-based ordinal regression network (CORNet), to reconstruct monocular depth maps in the ordinal regression manner with context information in this paper. Firstly, we put forward a novel context-based encoder with a feature transformation (FT) module to learn context information and details from inputs, and output multi-scale feature maps. Then, we design a boundary enhancement module (BEM) with a spatial attention mechanism following each operation of feature fusion, which captures boundary features in the scene to enhance the border depth. Finally, a feature optimization module (FOM) is designed to fuse and optimize the multi-scale features and boundary features to strengthen depth learning. What's more, we introduce an ordinal weighted inference to predict depth maps from probabilities and discretization values. Experiments and results on two challenging datasets, KITTI and NYU Depth V2, demonstrate that our proposed CORNet can estimate monocular depth maps effectively and obtain superior performance in capturing geometric features over existing methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
25  
被引频次合计:  
26  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:000819817700058  
Indexed Date:  
2022-07-10  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Elect Engn, Beijing Key Lab Work Safety Intelligent Monitorin, Beijing 100876, Peoples R China  
Institution Address:  
Univ Portsmouth, Sch Creat Technol, Portsmouth PO1 2DJ, Hants, England  
Institution:  
Institution:

### 293. AB-DLM: An Improved Deep Learning Model Based on Attention Mechanism and

摘要: Driver distraction behavior causes a large number of traffic accidents every year, resulting in economic losses and injuries. Currently, the driver still plays an important role in the driving and control of the vehicle due to the low level of vehicle automation and the immature development of autonomous driving. Therefore, it is vital to research distraction detection for drivers. However, in realistic driving scenarios with uncertain information, they are still some challenges in efficient and accurate driver distraction detection. In this paper, an improved deep learning model based on attention mechanisms and bi-directional feature pyramid networks (BiFPN) is proposed to identify driver distractions. Firstly, an improved data augmentation strategy is introduced to increase the data diversity to enhance the generalization capability of the model. Secondly, the squeeze-and-excitation (SE) attention mechanism layer is used after the C3 module of the original backbone network to enhance the important feature information and suppress the minor feature information. Finally, the BiFPN module is introduced into the neck network to better achieve multi-scale feature fusion without increasing the calculation amount too much. The experimental results show that the method proposed in this paper has an average mean accuracy rate (mAP) of 0.956 on the test set. Compared to the original model the mAP has improved by 13.2%. The detection speed of the model is 71 frames per second, and the memory occupation is 15.9 MB. This method has the advantages of high recognition accuracy, fast detection speed, and small memory occupation of the model, which are important for achieving engineering deployment.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
16  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
17  
ISSN:  
2169-3536  
入藏号:  
WOS:000842088500001  
Indexed Date:  
2022-08-28  
Institution Address:  
Lanzhou Jiaotong Univ, Sch Automat & Elect Engn, Lanzhou 730070, Peoples R China  
Institution Address:  
Shaanxi Kangfu Hosp, Childrens Rehabil Dept, Xian 710065, Peoples R China  
Institution:  
Institution:  
Shaanxi Kangfu Hosp

### 294. Multi-scale fusion for few-shot remote sensing image classification

摘要: Few-shot image classification is currently valuable research in the field of remote sensing image applications. The difficulty in obtaining training data for a large number of labelled remote sensing images leads to difficulties in applying traditional deep learning-based remote sensing image classification methods. Existing few-shot remote sensing image classification methods usually use a large number of base datasets to train network models in the pre-training phase and perform few-shot classification tasks after model fine-tuning on new datasets in the meta-testing phase. However, an insufficient number of remote sensing images lead to inaccurate features being extracted by the model, and the large amount of new class data for fine-tuning is also not easily accessible. In addition, using a pre-trained model to extract features from the new dataset will create a"negative transfer" problem. In this paper, we aim to address the above challenges in two ways. First, we use a subset of ImageNet, a readily available few-shot natural image dataset, for model pre-training, and eliminate the fine-tuning operation to simulate a real-world application scenario where remote sensing data is extremely scarce. Second, due to the significant differences in scale and style between the ImageNet dataset and the remote sensing dataset, which lead to a serious negative transfer problem in the model, we design a multi-scale feature fusion module to comprehensively decide the labels of query samples considering all scales to compensate for the scale differences and alleviate the"negative transfer" problem. We conducted experiments on four benchmark remote sensing datasets and achieved satisfactory performance.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:001074813000001  
Indexed Date:  
2023-10-19  
Institution Address:  
China Univ Petr East China, Coll Control Sci & Engn, Qingdao, Shandong, Peoples R China  
Institution Address:  
China Univ Petr East China, Coll Oceanog & Space Informat, Dongying, Shandong, Peoples R China  
Institution Address:  
China Univ Petr East China, Coll Control Sci & Engn, Qingdao 266580, Shandong, Peoples R China  
Institution:  
Institution:  
Institution:

### 295. PMENet: a parallel UNet based on the fusion of multiple attention

摘要: The presence of road cracks significantly impacts both traffic safety and road maintenance. Therefore, accurate detection of road cracks plays a crucial role in road maintenance and management. This study focused on addressing the critical challenge of accurate road crack detection in images through the development of a novel network architecture (PMENet) based on deep learning semantic segmentation algorithms. The proposed PMENet algorithm combined parallel contextual squeeze and excitation(PCSE), multiscale feature fusion(MFF) and double residual efficient attention(DREA) modules based on the UNet structure, which improves the ability to extract global local information and crack boundary information from the model. We thoroughly evaluate the performance of the PMENet algorithm on the publicly available Crack500 dataset by the ablation experiment. We compared our method with existing approaches, demonstrating its superiority in terms of accuracy, robustness, and generalization ability. The PMENet algorithm achieves an accuracy of 96.14%, an F1-score of 84.22%, and an IoU of 62.00%. The proposed model achieves a significant improvement of 5.9% in F1-score and 5.49% in IoU compared to the UNet model.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001215340900003  
Indexed Date:  
2024-05-12  
Institution Address:  
Hangzhou Dianzi Univ, Sch Mech Engn, Hangzhou 310018, Peoples R China  
Institution Address:  
Zhejiang Univ, Coll Elect Engn, Hangzhou 310027, Peoples R China  
Institution Address:  
Quzhou Univ, Coll Mech Engn, Quzhou 324000, Peoples R China  
Institution:  
Institution:  
Institution:

### 296. Radial basis function networks with lightweight multiscale fusion

摘要: A novel underwater picture enhancement approach under non-uniform lighting is presented to solve the issues of underwater photographs with unevenness due to additional lighting in deep-sea and night-time environments. Water suspended particles can cause image noise, low contrast, and colour deviation. The heterogeneous feature fusion module aims to combine multiple levels and levels of features with improving the network's ability to perceive semantic and specific information. The capability of autonomous underwater and remotely driven cars to explore and comprehend their environments is contingent on improving underwater images, a crucial low-level computer vision challenge. Recent applications of deep learning models include enhancing aquatic image quality and resolving several computer vision problems. Although several deep learning-based techniques exist for enhancing underwater images, their implementation is challenging due to the high memory and model parameter requirements. We propose a solution based on radial basis function networks (RBFN) for lightweight multiscale data fusion (LMFS). The LMFS incorporates diverse branches with varying kernel sizes to generate multiscale feature maps. The proposed RBFN-LMFS The convolution layer with jump connection and the attention module produces the output from the feature extraction module, which aims to extract various features at the network's beginning. The outcomes of our experiments on diverse data sets demonstrate that our proposed RBFN-LMFS technique performs well in processing both synthetic and authentic underwater images and successfully recovers image colour and texture characteristics. The visual output is superior to existing underwater image enhancement algorithms and is consistent with aspects of human vision.  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0266-4720  
eISSN:  
1468-0394  
入藏号:  
WOS:001007199400001  
Indexed Date:  
2023-06-23  
Institution Address:  
SRM Inst Sci & Technol, Fac Engn & Technol, Dept Informat Technol, Chennai, India  
Institution Address:  
Bennett Univ, Dept Comp Sci & Engn, Greater Noida, India  
Institution Address:  
Univ Cumberlands, Dept Informat Technol, Williamsburg, KY USA  
Institution Address:  
Koneru Lakshmaiah Educ Fdn, Dept Elect & Commun & Engn, Guntur, India  
Institution:  
Institution:  
Bennett Univ  
Institution:  
Institution:

### 297. 被撤回的出版物: Research on Land Utilization Spatial Classification Planning

摘要: With the development of China's social economy as well as the accelerating urbanization construction and the expanding scale of cities, the integration of land use and urban land classification based on land use spatial planning has become an important task for the sustainable development of China at present. Land use spatial classification planning is the basic basis for all kinds of development and protection construction activities, and government land use spatial planning at all levels plays an important role in implementing major national, provincial, and municipal strategies and promoting the rational and effective use of land use space. By briefly describing the spatial classification of land use and analyzing the idea of systematic integration of land use, this paper provides guidance and reference for exploring the construction of urban land use classification under land use spatial planning, aiming to improve the classification system of land use spatial planning. A neural network-based land use classification algorithm is proposed for the problems of few labeled samples of remote sensing images with high spatial resolution and feature deformation due to sensor height changes in land use spatial classification planning. By multiscale adaptive fusion of multiple convolutional layer features, the impact of feature deformation on classification accuracy is reduced. To further improve the classification accuracy, the depth features extracted from the pretraining network are used to pretrain the multiscale feature fusion part and the fully connected layer, and the whole network is fine-tuned using the augmented dataset. The experimental results show that the adaptive fusion method improves the fusion effect and effectively improves the accuracy of land use spatial classification planning.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1748-670X  
eISSN:  
1748-6718  
入藏号:  
WOS:000861389100002  
PubMed ID:  
36072770  
Indexed Date:  
2022-10-07  
Institution Address:  
Nanjing Univ, Sch Geog & Ocean Sci, Nanjing 210023, Peoples R China  
Institution Address:  
Jiangsu Prov Land Survey & Planning Inst, Nanjing 210093, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Geog Sci, Nanjing 210044, Peoples R China  
Institution:  
Institution:  
Jiangsu Prov Land Survey & Planning Inst  
Institution:

### 298. Remote Sensing Scene Classification Based on Attention-Enabled

摘要: Remote sensing image scene classification plays a significant role in remote sensing image analysis. Aiming at the problems of large transformation and scale variation of background and key objects in remote sensing images, we propose a neural architecture search (NAS) method based on attention search space. The network adaptively searches convolution, pooling, and attention operations in the appropriate layers. To ensure the stability of the searching process, a multistage network progressive fusion search method is proposed, which discards useless operations in stages, reduces the burden of search algorithm, and improves the search efficiency. Finally, paying attention to the association information between objects and scenes, a bottom-up multiscale fusion network connection strategy is proposed to fully reuse the semantics of multiscale feature maps in each stage. The experimental results show that the proposed method performs better than the manual method and the current neural network architecture search method.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
12  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
12  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000838694300018  
Indexed Date:  
2022-08-25  
Institution Address:  
Northwestern Polytech Univ, Unmanned Syst Res Inst, Xian 710000, Peoples R China  
Institution Address:  
Alibaba Grp, Beijing 100015, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Sch Artificial Intelligence Opt & Elect iOPEN, Xian 710000, Peoples R China  
Institution:  
Institution:  
Institution:

### 299. A Deep Learning-Based Buffalo Optimizer based Squeeze and Excitation

摘要: A Squeeze and Excitation Network is a deep -learning architectural component designed to enhance networks. The "squeeze" step reduces the spatial dimensions of the input feature maps, and the "excitation" step adaptively recalibrates channel -wise feature responses. This allows the network to focus on more educational features and ignore less useful ones. Garbage classification is a crucial task for sustainable environmental management. It involves categorizing waste into recyclables, organics, and non-recyclables, among other classes. Deep learning models, like the proposed Buffalo Optimizer -based SEN, can play a pivotal role in automating this classification process using computer vision techniques. Garbage Classification: There is still a long way to go until countries worldwide have successfully increased public awareness and implemented measures to prevent the rapid degradation of the natural environment. The annual global generation of e -waste is between 20 and 50 million growing components of municipal solid garbage. The disposal of electronic trash presents significant threats to environmental quality. As a result, pollution monitoring and control are crucial for maintaining a healthy ecosystem. In this research, picture -layered characteristics were extracted using an encoder composed of varying numbers of ResBlocks and the Squeeze -and -Excitation (SE) block, which was built on top of the UNet backbone network. UNet's decoder structure was streamlined, and the number of network model parameters was cut in half. In the meantime, the multiscale feature fusion module was developed to enhance the network's detection accuracy by optimizing its parameters with a bespoke loss function in place of the standard function. The African Buffalo Optimisation Algorithm is also used to fine-tune the hyperparameters.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1790-7632  
入藏号:  
WOS:001245949500023  
Indexed Date:  
2024-11-03  
Institution Address:  
Sona Coll Technol, Dept Informat Technol, Salem, India  
Institution Address:  
Univ Coll Engn, Dept Elect & Commun Engn, Panruti 607106, India  
Institution Address:  
Panimalar Engn Coll, Dept Informat Technol, Chennai, India  
Institution Address:  
Anna Univ, BIT Campus, Univ Coll Engn, Dept Elect & Commun Engn, Trichy, Tamil Nadu, India  
Institution:  
Institution:  
Univ Coll Engn  
Institution:  
Panimalar Engn Coll  
Institution:

### 300. Multi-Scale Inter-Communication Spatio-Temporal Network for Video

摘要: Video compression artifacts are widespread in online videos, which greatly affects the quality of the videos. In recent years, the deep learning-based methods for video compression artifacts reduction have made impressive achievements. However, most of existing methods ignore the rich multi-scale information contained in video frames. In this brief, we propose a new multi-scale inter-communication spatio-temporal network (MSICSTN) for video compression artifacts reduction by fully exploring contextual multi-scale information. MSICSTN is constructed on the basis of the multi-scale feature inter-communication module (MSFICM) and the source feature selection enhancement module (SFSEM). Specifically, the MSFICM adaptively generates respective attention weights using channel statistical information after the inter-communication of different scale features, and recalibrates the input multi-scale features to enhance the multi-scale features. The SFSEM utilizes the statistical information of the shallow features of the target frame to fuse the shallow features with the deep features for better dense pixel predictions. The experimental results on test videos demonstrate that our method outperforms the existing methods.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
1549-7747  
eISSN:  
1558-3791  
入藏号:  
WOS:000965335000001  
Indexed Date:  
2023-05-31  
Institution Address:  
Sichuan Univ, Coll Elect & Informat Engn, Chengdu 610065, Peoples R China  
Institution:

### 301. Imaging Segmentation of Brain Tumors Based on the Modified U-net Method

摘要: Brain tumor segmentation in medical image analysis is a challenging task. Deep learning techniques have recently shown promise in resolving a variety of computer vision problems, such as semantic segmentation and image classification. Brain MRI (magnetic resonance imaging) requires precise brain image segmentation for effective, rapid diagnosis and treatment planning. However, it is quite difficult to manually segment the brain image rapidly and accurately in low-quality, noisy images. This paper proposes a U-Net and combined attention mechanism-based method. This research enhances the segmentation of images of tumors in the brain using modified U-net. Traditional U-net segmentation techniques are still widely used in the medical field, but they have a number of limitations when dealing with small targets and fuzzier boundaries. To address this issue, we made the following modifications to U-net: We propose attention mechanisms to assist the network in concen trating on important regions. The multiscale feature fusion strategy improves the efficacy of network segmen- tation at various scales. Cross-entropy loss function and data augmentation improve the performance of the network further. Our method was validated usingthe Brats2019 dataset. The experimental results demonstrate that our proposed methodology exhibits superior speed and efficiency compared to existing techniques in the context ofbrain image segmentation. The dice coefficients for the multiple branch TS-U-net model were 0.876, 0.868, and 0.814 in the tumor subregions of WT, TC, and ET, respectively. This exemplifies the feasibility and potential of our methodology for the segmentation of medical images.  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1392-124X  
入藏号:  
WOS:001384432900007  
Indexed Date:  
2025-01-03  
Institution Address:  
Wenzhou Polytech, Sch Artificial Intelligence, Wenzhou 325000, Zhejiang, Peoples R China  
Institution Address:  
Univ Teknikal Malaysia Melaka, Fac Informat & Commun Technol, Melaka, Malaysia  
Institution Address:  
Wenzhou Med Univ, Wenzhou Cent Hosp, Dingli Clin Inst, Div Nephrol, Wenzhou 325000, Zhejiang, Peoples R China  
Institution:  
Institution:  
Institution:

### 302. 6mAPred-MSFF: A Deep Learning Model for Predicting DNA N6-Methyladenine

摘要: DNA methylation is one of the most extensive epigenetic modifications. DNA N6-methyladenine (6mA) plays a key role in many biology regulation processes. An accurate and reliable genome-wide identification of 6mA sites is crucial for systematically understanding its biological functions. Some machine learning tools can identify 6mA sites, but their limited prediction accuracy and lack of robustness limit their usability in epigenetic studies, which implies the great need of developing new computational methods for this problem. In this paper, we developed a novel computational predictor, namely the 6mAPred-MSFF, which is a deep learning framework based on a multi-scale feature fusion mechanism to identify 6mA sites across different species. In the predictor, we integrate the inverted residual block and multi-scale attention mechanism to build lightweight and deep neural networks. As compared to existing predictors using traditional machine learning, our deep learning framework needs no prior knowledge of 6mA or manually crafted sequence features and sufficiently capture better characteristics of 6mA sites. By benchmarking comparison, our deep learning method outperforms the state-of-the-art methods on the 5-fold cross-validation test on the seven datasets of six species, demonstrating that the proposed 6mAPred-MSFF is more effective and generic. Specifically, our proposed 6mAPred-MSFF gives the sensitivity and specificity of the 5-fold cross-validation on the 6mA-rice-Lv dataset as 97.88% and 94.64%, respectively. Our model trained with the rice data predicts well the 6mA sites of other five species: Arabidopsis thaliana, Fragaria vesca, Rosa chinensis, Homo sapiens, and Drosophila melanogaster with a prediction accuracy 98.51%, 93.02%, and 91.53%, respectively. Moreover, via experimental comparison, we explored performance impact by training and testing our proposed model under different encoding schemes and feature descriptors.  
在 BIOSIS Citation Index 中的被引频次:  
7  
Web of Science 核心合集中的 "被引频次":  
8  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
eISSN:  
2076-3417  
入藏号:  
WOS:000688620100001  
Indexed Date:  
2021-09-02  
Institution Address:  
Xiamen Univ, Sch Informat, Dept Software Engn, Xiamen 361005, Peoples R China  
Institution:

### 303. Improved CLIP-ILP Model for Detecting Illegal Passenger Transport in

摘要: In recent years, illegal passenger transport in freight trucks has become a critical concern for traffic safety and law enforcement. This study proposes an automated system for detecting illegal passenger transport using an improved CLIP-ILP (Illegal Passenger Detection) model. The proposed model incorporates multi-scale feature fusion, a cross-modal self-attention mechanism, and a more powerful text encoder to enhance detection performance. To evaluate the model, we constructed a comprehensive dataset consisting of tens of thousands of truck images, categorized into two types of trucks (four-wheeled and three-wheeled) and further classified into two subcategories: "illegal passenger transport" and "non-illegal passenger transport." The model was trained using this dataset, with an emphasis on leveraging the CLIP framework's ability to understand and integrate visual and textual data. Experimental results demonstrate that the proposed CLIP-ILP model achieves superior accuracy and robustness in detecting illegal passenger transport under various conditions. This research not only highlights the potential of deep learning technologies in enhancing traffic safety but also provides a novel and efficient approach for law enforcement agencies to monitor and address this growing issue effectively. You can access the code for our proposed method at https://github.com/wu-xuan-git/CLIP-ILP.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001405911400004  
Indexed Date:  
2025-02-01  
Institution Address:  
Yunnan Minzu Univ, Sch Elect & Informat Technol, Kunming 650500, Peoples R China  
Institution Address:  
Yunnan Minzu Univ, Yunnan Key Lab Unmanned Autonomous Syst, Kunming 650500, Peoples R China  
Institution Address:  
Yunnan Minzu Univ, Sch Math & Comp Sci, Kunming 650500, Peoples R China  
Institution:  
Institution:  
Institution:

### 304. Multi-scale feature fusion network for pixel-level pavement distress

摘要: Automatic pavement distress detection is essential to monitoring and maintaining pavement condition. Currently, many deep learning-based methods have been utilized in pavement distress detection. However, distress segmentation remains as a challenge under complex pavement conditions. In this paper, a novel deep neural network architecture, W-segnet, based on multi-scale feature fusions, is proposed for pixel-wise distress segmentation. The proposed W-segnet concatenates distress location information with distress classification features in two symmetric encoder-decoder structures. Three major types of distresses: crack, pothole, and patch are segmented and the results were discussed. Experimental results show that the proposed W-segnet is robust in various scenarios, achieving a mean pixel accuracy WPM of 87.52% and a mean intersection over union (MIoU) of 75.88%. The results demonstrate that W-segnet outperforms other state-of-the-art semantic segmentation models of U-net, SegNet, and PSPNet. Comparison of cost of model training and inference indicates that W-segnet has the largest number of parameters, which needs a slightly longer training time while it does not increase the inference cost. Four public datasets were used to test the generalization ability of the proposed model and the results demonstrate that the W-segnet possesses well segmentation performance.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
50  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
54  
ISSN:  
0926-5805  
eISSN:  
1872-7891  
入藏号:  
WOS:000833418700002  
Indexed Date:  
2022-12-15  
Institution Address:  
Southeast Univ, Sch Transportat, Nanjing 211189, Peoples R China  
Institution:

### 305. EchoEFNet: Multi-task deep learning network for automatic calculation of

摘要: Left ventricular ejection fraction (LVEF) is essential for evaluating left ventricular systolic function. However, its clinical calculation requires the physician to interactively segment the left ventricle and obtain the mitral annulus and apical landmarks. This process is poorly reproducible and error prone. In this study, we propose a multi-task deep learning network EchoEFNet. The network use ResNet50 with dilated convolution as the backbone to extract high-dimensional features while maintaining spatial features. The branching network used our designed multi-scale feature fusion decoder to segment the left ventricle and detect landmarks simultaneously. The LVEF was then calculated automatically and accurately using the biplane Simpson's method. The model was tested for performance on the public dataset CAMUS and private dataset CMUEcho. The experimental results showed that the geometrical metrics and percentage of correct keypoints of EchoEFNet outperformed other deep learning methods. The correlation between the predicted LVEF and true values on the CAMUS and CMUEcho datasets was 0.854 and 0.916, respectively.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
18  
被引频次合计:  
19  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:000949460900001  
PubMed ID:  
36863190  
Indexed Date:  
2023-03-30  
Institution Address:  
Northeastern Univ, Comp Sci & Engn, Shenyang, Peoples R China  
Institution Address:  
Northeastern Univ, Key Lab Intelligent Comp Med Image, Minist Educ, Shenyang, Peoples R China  
Institution Address:  
China Med Univ, Dept Cardiovasc Ultrasound, Hosp 1, Shenyang, Peoples R China  
Institution:  
Institution:  
Institution:

### 306. Learning conditional photometric stereo with high-resolution features

摘要: Photometric stereo aims to reconstruct 3D geometry by recovering the dense surface orientation of a 3D object from multiple images under differing illumination. Traditional methods normally adopt simplified reflectance models to make the surface orientation computable. However, the real reflectances of surfaces greatly limit applicability of such methods to real-world objects. While deep neural networks have been employed to handle non-Lambertian surfaces, these methods are subject to blurring and errors, especially in high-frequency regions (such as crinkles and edges), caused by spectral bias: neural networks favor low-frequency representations so exhibit a bias towards smooth functions. In this paper, therefore, we propose a self-learning conditional network with multi-scale features for photometric stereo, avoiding blurred reconstruction in such regions. Our explorations include: (i) a multi-scale feature fusion architecture, which keeps high-resolution representations and deep feature extraction, simultaneously, and (ii) an improved gradient-motivated conditionally parameterized convolution (GM-CondConv) in our photometric stereo network, with different combinations of convolution kernels for varying surfaces. Extensive experiments on public benchmark datasets show that our calibrated photometric stereo method outperforms the state-of-the-art.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
22  
被引频次合计:  
23  
ISSN:  
2096-0433  
eISSN:  
2096-0662  
入藏号:  
WOS:000712590200007  
Indexed Date:  
2021-11-17  
Institution Address:  
Ocean Univ China, Dept Comp Sci & Technol, Qingdao 266100, Peoples R China  
Institution Address:  
Peking Univ, Wangxuan Inst Comp Technol, Beijing 100871, Peoples R China  
Institution Address:  
Shandong Univ Finance & Econ, Sch Comp Sci & Technol, Jinan 250002, Peoples R China  
Institution:  
Institution:  
Institution:

### 307. MultiNet: A deep neural network approach for detecting breast cancer

摘要: Breast cancer diagnosis from biopsy tissue images conducted manually by pathologists is costly, time-consuming, and disagreements among specialists. Nowadays, the advancement of the Computer-Aided Diagnosis (CAD) system allows pathologists to identify breast cancer more reliably and quickly.For this reason, interest in CAD-based deep learning models has been increased significantly. In this study, we propose a "MultiNet" framework based on the transfer learning concept to classify different breast cancer types using two publicly available datasets that include 7909 and 400 microscopic breast images, respec-tively. The proposed "MultiNet" framework is designed to provide fast and accurate diagnostics for breast cancer with binary classification (benign and malignant) and multi-class classification (benign, in situ, invasive, and normal). In the proposed framework, features from microscopy images are extracted using three well-known pre-trained models, including DenseNet-201, NasNetMobile, and VGG16. The extracted features are then fed into the concatenate layer, making a robust hybrid model. The proposed framework yields an overall classification accuracy of 99% in classifying two classes. It also achieves 98% classification accuracy in classifying four classes. Such promising results will provide the opportunity to use "MultiNet" framework as a diagnostic model in clinics and health care.(c) 2021 The Authors. Published by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).  
在 BIOSIS Citation Index 中的被引频次:  
4  
Web of Science 核心合集中的 "被引频次":  
39  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
41  
ISSN:  
1319-1578  
eISSN:  
2213-1248  
入藏号:  
WOS:000862928800002  
Indexed Date:  
2022-10-15  
Institution Address:  
Mawlana Bhashani Sci & Technol Univ, Dept Comp Sci & Engn, Tangail, Bangladesh  
Institution Address:  
Natl Inst Text Engn & Res, Dept Comp Sci & Engn, Dhaka, Bangladesh  
Institution:  
Institution:  
Natl Inst Text Engn & Res

### 308. BMFNet: Bifurcated multi-modal fusion network for RGB-D salient object

摘要: Although deep learning-based RGB-D salient object detection methods have achieved impressive results in the recent years, there are still some issues need to be addressed including multi-modal fusion and multi-level aggregation. In this paper, we propose a bifurcated multi-modal fusion network (BMFNet) to address these two issues cooperatively. First, we design a multi-modal feature interaction (MFI) module to fully capture the complementary information between the RGB and depth features by leveraging the channel attention and spatial attention. Second, unlike the widely used layer-by-layer progressive fusion, we adopt a bifurcated fusion strategy for all the multi-level unimodal and cross-modal features to effectively reduce the gaps between features at different levels. For the intra-group feature aggregation, a multi-modal feature fusion (MFF) module is designed to integrate the intra-group multi-modal features to produce a low-level/high-level saliency feature. For the inter-group aggregation, a multi-scale feature learning (MFL) module is introduced to exploit the contextual interactions between different scales to boost fusion performance. Experimental results on five public RGB-D datasets demonstrate the effectiveness and superiority of our proposed network. The code and prediction maps will be available at https://github.com/ZhangQing0329/BMFNet  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
0262-8856  
eISSN:  
1872-8138  
入藏号:  
WOS:001240162600001  
Indexed Date:  
2024-06-11  
Institution Address:  
Shanghai Inst Technol, Sch Comp Sci & Informat Engn, Shanghai 201418, Peoples R China  
Institution Address:  
Shanghai Inst Technol, Sch Mech Engn, Shanghai 201418, Peoples R China  
Institution:  
Institution:

### 309. Lightweight Semantic Segmentation Neural Network for Autonomous Driving

摘要: Image semantic segmentation has very important applications in autonomous driving, allowing robots to segment semantic information in the environment to make decisions about downstream control actions. However, most of the deep learning models for this task are relatively large, require huge computing resources, and are difficult to use in mobile devices. In order to solve this problem, a lightweight neural network model for semantic segmentation is proposed, which uses a network architecture combining encoding-decoding and two-branch type. Grouping convolution, deep separable convolution, multi-scale feature fusion module and channel shuffling technology are used to reduce the number of network parameters and improve the prediction accuracy of the model. The model training in this paper combines Adam training method and stochastic gradient descent method. The Cityscapes data set is used, and 1000 training cycles are set. After testing, the number of model parameters is 3.5\*10~6, and the calculation speed on a single graphics card GTX 1070Ti is 103 frames per second, which meets the real-time calculation standard. In the model evaluation indicators, the average intersection ratio is 61.3%, and the pixel accuracy rate is 93.4%, both of which are better than SegNet and ENet models.  
摘要:  
图像语义分割在自动驾驶领域有十分重要的应用,可以让机器人在环境中分割出语义信息,从而对下游的控制动作做出决策。但大部分的深度学习模型都比较大,需庞大的计算资源,很难在移动设备中使用。为了解决这个问题,提出了一种用于语义分割的轻量级神经网络模型,采用编码-解码型与二分支型相结合的网络架构,利用分组卷积、深度可分离卷积、多尺度特征融合模块与通道混洗技术减少网络参数量,提升模型预测精度。该模型训练结合Adam训练法与随机梯度下降法,使用Cityscapes数据集,设置1 000个训练周期。经测试,该模型参数量为3.5\*10~6,在单张显卡Nvidia GTX 1070Ti上的运算速度为每秒103帧图片,达到实时计算标准。在模型评估指标中,平均交并比为61.3%,像素准确率为93.4%,性能均优于SegNet和ENet两种模型。  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
1  
ISSN:  
1002-8331  
入藏号:  
CSCD:7494752  
Indexed Date:  
2023-09-28  
Institution Address:  
广东海洋大学数学与计算机学院, 湛江, 广东 524088, 中国  
Institution Address:  
广东海洋大学电子与信息工程学院, 湛江, 广东 524088, 中国  
Institution Address:  
School of Mathematics and Computer, Guangdong Ocean University, Zhanjiang, Guangdong 524088, China  
Institution Address:  
School of Electronics and Information Engineering, Guangdong Ocean University, Zhanjiang, Guangdong 524088, China  
Institution:  
广东海洋大学数学与计算机学院  
Institution:  
广东海洋大学电子与信息工程学院  
Institution:  
School of Mathematics and Computer, Guangdong Ocean University  
Institution:  
School of Electronics and Information Engineering, Guangdong Ocean University

### 310. Enhanced You Only Look Once X for surface defect detection of strip

摘要: Using deep learning-based methods to detect surface defects in strip steel can reduce the impact of human factors and lower costs while maintaining accuracy and efficiency. However, the main disadvantages of this method is the inability to tradeoff accuracy and efficiency. In addition, the low proportion of valid information and the lack of distinctive features result in a high rate of missed detection of small objects. In this paper, we propose a lightweight YOLOX surface defect detection network and introduce the Multi-scale Feature Fusion Attention Module (MFFAM). Lightweight CSP structures are used to optimize the backbone of the original network. MFFAM uses different scales of receptive fields for feature maps of different resolutions, after which features are fused and passed into the spatial and channel attention modules in parallel. Experimental results show that lightweight CSP structures can improve the detection frame rate without compromising accuracy. MFFAM can significantly improve the detection accuracy of small objects. Compared with the initial YOLOX, the mAP and FPS were 81.21% and 82.87Hz, respectively, which was an improvement of 4.29% and 12.72Hz. Compared with existing methods, the proposed model has superior performance and practicality, verifying the effectiveness of the optimization method.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1662-5218  
入藏号:  
WOS:000894201700001  
PubMed ID:  
36479529  
Indexed Date:  
2022-12-18  
Institution Address:  
Yancheng Inst Technol, Sch Informat Technol, Yancheng, Peoples R China  
Institution Address:  
Yancheng Xiongying Precis Machinery Co Ltd, Yancheng, Peoples R China  
Institution:  
Institution:  
Yancheng Xiongying Precis Machinery Co Ltd

### 311. EKFNet: edge-based Kalman filter network for real-time EEG signal

摘要: Objective. Signal denoising methods based on deep learning have been extensively adopted on electroencephalogram devices. However, they are unable to deploy on edge-based portable or wearable (P/W) electronics due to the high computational complexity of the existed models. To overcome such issue, we propose an edge-based lightweight Kalman filter network (EKFNet) that does not require manual prior knowledge estimation. Approach. Specifically, we construct a multi-scale feature fusion module to capture multi-scale feature information and implicitly compute the prior knowledge. Meanwhile, we design an adaptive gain estimation module that incorporates long short-term memory and sequential channel attention module to dynamically predict the Kalman gain. Furthermore, we present an optimization strategy utilizing operator fusion and constant folding to reduce the model's computational overhead and memory footprint. Main results. Experimental results show that the EKFNet reduces the sum of the square of the distances by at least 12% and improves the cosine similarity by at least 2.2% over the state-of-the-art methods. Besides, the model optimization shortens the inference time by approximately 3.3x. The code of our EKFNet is available at https://github.com/cathnat/EKFNet. Significance. By integrating Kalman filter with deep learning, the approach addresses the parameter-setting challenges in traditional algorithms while reducing computational overhead and memory consumption, which exhibits a good tradeoff between algorithm performance and computing power.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1741-2560  
eISSN:  
1741-2552  
入藏号:  
WOS:001380298800001  
PubMed ID:  
39622170  
Indexed Date:  
2024-12-25  
Institution Address:  
Minjiang Univ, Sch Comp & Big Data, Fujian Prov Key Lab Informat Proc & Intelligent Co, Fuzhou 350121, Peoples R China  
Institution Address:  
Fuzhou Univ, Coll Comp & Big Data, Fuzhou 350108, Peoples R China  
Institution Address:  
Univ Lahore, Dept Comp Engn, Lahore 54000, Pakistan  
Institution Address:  
Shenzhen Univ, Coll Elect & Informat Engn, Shenzhen 518060, Peoples R China  
Institution Address:  
Univ Politecn Cataluna, Dept Automat Control Tech, Barcelona 08034, Spain  
Institution Address:  
Putian Univ, New Engn Ind Coll, Putian 351100, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 312. CDPNet: conformer-based dual path joint modeling network for bird sound

摘要: Bird species monitoring is important for the preservation of biological diversity because it provides fundamental information for biodiversity assessment and protection. Automatic acoustic recognition is considered to be an essential technology for realizing automatic monitoring of bird species. Current deep learning-based bird sound recognition methods do not fully conduct long-term correlation modeling along both the time and frequency axes of the spectrogram. Additionally, these methods have not completely studied the impact of different scales of features on the final recognition. To solve the abovementioned problems, we propose a Conformer-based dual path joint modeling network (CDPNet) for bird sound recognition. To the best of our knowledge, this is the first attempt to adopt Conformer in the bird sound recognition task. Specifically, the proposed CDPNet mainly consists of a dual-path time-frequency joint modeling module (DPTFM) and a multi-scale feature fusion module (MSFFM). The former aims to simultaneously capture time-frequency local features, long-term time dependence, and long-term frequency dependence to better model bird sound characteristics effectively. The latter is designed to improve recognition accuracy by fusing different scales of features. The proposed algorithm is implemented on an edge computing platform, NVIDIA Jetson Nano, to build a real-time bird sound recognition monitoring system. The ablation experimental results verify the benefit of using the DPTFM and the MSFFM. Through training and testing on the Semibirdaudio dataset containing 27,155 sound clips and the public Birdsdata dataset, the proposed CDPNet outperforms the other state-of-the-art models in terms of F1-score, precision, recall, and accuracy.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:001171681100002  
Indexed Date:  
2024-03-13  
Institution Address:  
Chinese Acad Sci, Inst Semicond, Lab Solid State Optoelect Informat Technol, Beijing 100083, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution Address:  
Shandong Normal Univ, Jinan 250014, Shandong, Peoples R China  
Institution:  
Institution:  
Institution:

### 313. Multi-scale feature fusion network with local attention for lung

摘要: Computer-assisted medical care can benefit from the lung region segmentation method. Numerous methods provide end-to-end solutions, these methods employ convolution neural networks to segment lung regions from images. The low contrast, unpredictable appearance, and other problems in medical images have an effect on the accuracy of existing methods. In order to overcome the aforementioned issues, the MSDC (multi-scale dilated convolution) module is added to the short-cut connection, so as to fuse multi-scale features with various receptive fields to obtain more global information of lung area. Moreover, a local attention module which includes channel attention and spatial attention is suggested to give more weight to the lung area to lower the influence of background. Several lung segmentation datasets are employed to evaluate the segmentation performance of images qualitatively and quantitatively. From the experimental results, we can see that the segmentation accuracy of our model outperforms many recent image segmentation methods.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0923-5965  
eISSN:  
1879-2677  
入藏号:  
WOS:001074391000001  
Indexed Date:  
2023-10-08  
Institution Address:  
Hebei Univ Sci & Technol, Dept Pharm, Shijiazhuang, Peoples R China  
Institution Address:  
Shijiazhuang Hosp Tradit Chinese Med, Dept Radiol, Shijiazhuang, Peoples R China  
Institution Address:  
Hebei Gen Hosp, Dept Ultrasound, Shijiazhuang, Peoples R China  
Institution:  
Institution:  
Shijiazhuang Hosp Tradit Chinese Med  
Institution:  
Hebei Gen Hosp

### 314. MDOAU-Net: A Lightweight and Robust Deep Learning Model for SAR Image

摘要: Offshore aquaculture raft information extraction from synthetic aperture radar (SAR) images is essential for large-scale marine resource exploitation and protection. In this letter, a deep learning model called multi-scaled attention U-net with dilated convolution and offset convolution (MDOAU-net) is proposed for aquaculture raft monitoring via SAR image segmentation. The U-net backbone and attention gate of the Attention U-net are used in the MDOAU-net model. In addition, the MDOAU-net model consists of three distinctive parts. First, a multi-scale feature-fusion block is adopted in its input to extract features from raw images. Moreover, adapted from the Attention U-net for SAR image segmentation, fewer channels are used in each convolution layer of the MDOAU-net to match latent features in SAR images. Furthermore, nine dilated convolution blocks are adopted in the encoder-decoder structure to extract semantic features in the presence of speckle noises. In addition, offset convolution blocks are developed to convert spatial information into channel information for the precise segmentation of blurry boundaries. Four skip connections of the U-net backbone are replaced by four offset convolution blocks. Experimental results are elaborated to demonstrate the superior performance of the MDOAU-net model to seven existing methods in terms of overall accuracy (OA) and number of parameters.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
27  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
27  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000756792900002  
Indexed Date:  
2022-03-04  
Institution Address:  
City Univ Hong Kong, Sch Data Sci, Hong Kong, Peoples R China  
Institution Address:  
Natl Marine Environm Monitoring Ctr, Dept Ocean Remote Sensing, State Environm Protect Key Lab Coastal Ecosyst, Dalian 116023, Peoples R China  
Institution Address:  
City Univ Hong Kong, Dept Comp Sci, Hong Kong, Peoples R China  
Institution:  
Institution:  
Institution:

### 315. Speech bandwidth extension based on multi-scale feature fusion

摘要: A new type of end-to-end neural network model was proposed by us,with the aim of addressing the issues of insufficient utilization of data features,long training periods,and low quality of generated speech by existing deep learning models. The integration of features from different data dimensions within the network model was encouraged,thereby reducing the utilization of data features and obtaining more low- and high-frequency mapping relationship solutions,resulting in a shortened overall training cycle of the model.To maximize the utilization of data features and enhance the weight ratio of key features in long time series data, a residual multi-head self-attention mechanism was designed. Additionally,a hybrid loss function based on the time-frequency domain and Mel spectrum was proposed to optimize the model.The wideband speech reconstructed by this method was evaluated subjectively and objectively,and the experimental results indicate that it is superior to traditional methods and some recent neural network-based speech bandwidth extension methods.  
摘要:  
针对现有的深度学习模型在语音频带扩展领域数据特征利用不充分、训练周期长以及生成语音质量不高等问题,提出了一种新型的端到端神经网络模型,该模型通过融合不同数据维度特征促使网络模型利用更少的数据特征量,获取更多的低高频映射关系解,从而减少模型的整体训练周期.为了提高长时序数据中关键特征的权重占比,设计了一种残差多头自注意力机制,从而达到数据特征利用率的最大化.此外,提出了一种基于时频域和Mel频谱的混合损失函数对模型进行优化.实验结果表明:该方法重构的宽带语音在主客观的评价中均优于传统方法和近年来的一些基于神经网络的语音频带扩展方法.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1671-4512  
入藏号:  
CSCD:7591240  
Indexed Date:  
2024-01-19  
Institution Address:  
江西理工大学信息工程学院, 赣州, 江西 341000, 中国  
Institution Address:  
江西理工大学信息工程学院;;中国科学院大学电子电气与通信工程学院, ;;, 赣州;;, 江西;;北京 341000;;100049, 中国  
Institution Address:  
中国电子技术标准化研究院, 北京 100007, 中国  
Institution Address:  
School of Information Engineering, Jiangxi University of Science and Technology, Ganzhou, Jiangxi 341000, China  
Institution Address:  
School of Information Engineering, Jiangxi University of Science and Technology;;School of Electronic, Electronical, and Communication Engineering, University of Chinese Academy of Sciences, ;;, Ganzhou;;, Jiangxi;;Beijing 341000;;100049  
Institution Address:  
China Electronic Technology Standardization Institute, Beijing 100007, China  
Institution:  
江西理工大学信息工程学院  
Institution:  
江西理工大学信息工程学院;;中国科学院大学电子电气与通信工程学院  
Institution:  
中国电子技术标准化研究院  
Institution:  
School of Information Engineering, Jiangxi University of Science and Technology  
Institution:  
School of Information Engineering, Jiangxi University of Science and Technology;;School of Electronic, Electronical, and Communication Engineering, University of Chinese Academy of Sciences  
Institution:  
China Electronic Technology Standardization Institute

### 316. Diabetic Plantar Foot Segmentation in Active Thermography Using a

摘要: Pathological conditions in diabetic feet cause surface temperature variations, which can be captured quantitatively using infrared thermography. Thermal images captured during recovery of diabetic feet after active cooling may reveal richer information than those from passive thermography, but diseased foot regions may exhibit very small temperature differences compared with the surrounding area, complicating plantar foot segmentation in such cold-stressed active thermography. In this study, we investigate new plantar foot segmentation methods for thermal images obtained via cold-stressed active thermography without the complementary information from color or depth channels. To better deal with the temporal variations in thermal image contrast when planar feet are recovering from cold immersion, we propose an image pre-processing method using a two-stage adaptive gamma transform to alleviate the impact of such contrast variations. To improve upon existing deep neural networks for segmenting planar feet from cold-stressed infrared thermograms, a new deep neural network, the Plantar Foot Segmentation Network (PFSNet), is proposed to better extract foot contours. It combines the fundamental U-shaped network structure, a multi-scale feature extraction module, and a convolutional block attention module with a feature fusion network. The PFSNet, in combination with the two-stage adaptive gamma transform, outperforms multiple existing deep neural networks in plantar foot segmentation for single-channel infrared images from cold-stressed infrared thermography, achieving an accuracy of 97.3% and 95.4% as measured by Intersection over Union (IOU) and Dice Similarity Coefficient (DSC) respectively.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
eISSN:  
1424-8220  
入藏号:  
WOS:001089511600001  
PubMed ID:  
37896605  
Indexed Date:  
2023-11-09  
Institution Address:  
Shenzhen Univ, Coll Mechatron & Control Engn, Shenzhen 518000, Peoples R China  
Institution Address:  
Chongqing Normal Univ, Coll Comp & Informat Sci, Chongqing 401331, Peoples R China  
Institution Address:  
Southern Med Univ, Shunde Hosp, Foshan 528000, Peoples R China  
Institution Address:  
Chongqing Med Univ, Dept Rheumatol & Immunol, Affiliated Hosp 2, Chongqing 400010, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 317. Improved YOLOv3 Network for Insulator Detection in Aerial Images with

摘要: Automatic inspection of insulators from high-voltage transmission lines is of paramount importance to the safety and reliable operation of the power grid. Due to different size insulators and the complex background of aerial images, it is a difficult task to recognize insulators in aerial views. Most of the traditional image processing methods and machine learning methods cannot achieve sufficient performance for insulator detection when diverse background interference is present. In this study, a deep learning method-based on You Only Look Once (YOLO)-will be proposed, capable of detecting insulators from aerial images with complex backgrounds. Firstly, aerial images with common aerial scenes were collected by Unmanned Aerial Vehicle (UAV), and a novel insulator dataset was constructed. Secondly, to enhance feature reuse and propagation, on the basis of YOLOv3 and Dense-Blocks, the YOLOv3-dense network was utilized for insulator detection. To improve detection accuracy for different sized insulators, a structure of multiscale feature fusion was adapted to the YOLOv3-dense network. To obtain abundant semantic information of upper and lower layers, multilevel feature mapping modules were employed across the YOLOv3-dense network. Finally, the YOLOv3-dense network and compared networks were trained and tested on the testing set. The average precision of YOLOv3-dense, YOLOv3, and YOLOv2 were 94.47%, 90.31%, and 83.43%, respectively. Experimental results and analysis validate the claim that the proposed YOLOv3-dense network achieves good performance in the detection of different size insulators amid diverse background interference.  
Web of Science 核心合集中的 "被引频次":  
57  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
61  
eISSN:  
2079-9292  
入藏号:  
WOS:000638362400001  
Indexed Date:  
2021-05-06  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Elect & Informat Engn, Nanjing 211106, Peoples R China  
Institution Address:  
Chizhou Univ, Coll Mech & Elect Engn, Chizhou 247000, Peoples R China  
Institution:  
Institution:

### 318. 被撤回的出版物: Ancient Stone Inscription Image Denoising and Inpainting

摘要: Chinese ancient stone inscriptions contain Chinese traditional calligraphy culture and art information. However, due to the long history of the ancient stone inscriptions, natural erosion, and poor early protection measures, there are a lot of noise in the existing ancient stone inscriptions, which has adverse effects on reading these stone inscriptions and their aesthetic appreciation. At present, digital technologies have played important roles in the protection of cultural relics. For ancient stone inscriptions, we should obtain more perfect digital results without multiple types of noise, while there are few deep learning methods designed for processing stone inscription images. Therefore, we propose a basic framework for image denoising and inpainting of stone inscriptions based on deep learning methods. Firstly, we collect as many images of stone inscriptions as possible and preprocess these images to establish an inscriptions image dataset for image denoising and inpainting. In addition, an improved GAN with a denoiser is used for generating more virtual stone inscription images to expand the dataset. On the basis of these collected and generated images, we designed a stone inscription image denoising model based on multiscale feature fusion and introduced Charbonnier loss function to improve this image denoising model. To further improve the denoising results, an image inpainting model with the coherent semantic attention mechanism is introduced to recover some effective information removed by the former denoising model as much as possible. The experimental results show that our image denoising model achieves better results on PSNR, SSIM, and CEI. The final results have obvious visual improvement compared with the original stone inscription images.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1026-0226  
eISSN:  
1607-887X  
入藏号:  
WOS:000797416500002  
Indexed Date:  
2022-05-29  
Institution Address:  
Northwest Univ, Sch Informat Sci & Technol, Xian, Peoples R China  
Institution Address:  
Northwest Univ, Art Sch, Xian, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Teaching Affairs Dept, Xian, Peoples R China  
Institution Address:  
Lvliang Univ, Dept Comp Sci & Technol, Lvliang 033000, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 319. PLU-Net: Extraction of multiscale feature fusion

摘要: In recent years, deep learning algorithms have achieved remarkable results in medical image segmentation. These networks with an enormous number of parameters often encounter challenges in handling image boundaries and details, which may result in suboptimal segmentation results. To solve the problem, we develop atrous spatial pyramid pooling (ASPP) and combine it with the squeeze-and-excitation block (SE block), as well as present the PS module, which employs a broader and multiscale receptive field at the network's bottom to obtain more detailed semantic information. We also propose the local guided block (LG block) and also its combination with the SE block to form the LS block, which can obtain more abundant local features in the feature map, so that more edge information can be retained in each down sampling process, thereby improving the performance of boundary segmentation. We propose PLU-Net and integrate our PS module and LS block into U-Net. We put our PLU-Net to the test on three benchmark datasets, and the results show that by fewer parameters and FLOPs, it outperforms on medical semantic segmentation tasks.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0094-2405  
eISSN:  
2473-4209  
入藏号:  
WOS:001108874500001  
PubMed ID:  
38010657  
Indexed Date:  
2023-12-14  
Institution Address:  
Beihang Univ, Sch Comp Sci & Engn, Beijing, Peoples R China  
Institution Address:  
Carnegie Mellon Univ, Robot Inst, Pittsburgh, PA USA  
Institution Address:  
Nankai Univ, Coll Artificial Intelligence, Tianjin, Peoples R China  
Institution Address:  
Beijing Univ Aeronaut & Astronaut, New Main Bldg,37 Xueyuan Rd,Huayuan Rd, Beijing 100191, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 320. Segmentation of breast nodules in ultrasound images based on multi-scale

摘要: Accurate breast nodule segmentation in ultrasound images is very challenging due to the low resolution and noise of ultrasound imaging, as well as the complexity and variability of the shape and texture of nodules, therefore, an endto- end automatic segmentation method of breast nodules in ultrasound images based on multi-scale and cross-spatial feature fusion was proposed. Firstly, a Multi-scale Feature Extraction and Fusion (MFEF) module was designed to enable the network to have multi-scale feature extraction ability by fusing four convolutional paths with different receptive fields. Then, for the multi-scale observation and information filtering of high-level semantic information, a Scale-aware Feature Aggregation (SFA) module was used at the bottleneck layer to enhance the deep feature extraction ability in the encoding stage. Besides, a Cross-spatial Residual Fusion (CRF) module was designed and applied to the skip connection between the encoder and decoder to fuse information among different encoding layers in a cross-spatial way and implement information complementarity between different encoding layers, further extract information features of encoding layer and narrow the difference between peer layers of encoder and decoder to better compensate for the information loss in the decoding stage. Experimental results on a public ultrasound breast nodule dataset show that the proposed method achieves DICE coefficient of 0.888, which is 0.033 to 0.094 higher than those of the mainstream deep learning segmentation models UNet, AttUNet, ResUNet++ and SKUNet, and is 0.001 to 0.068 higher than those of the improved models such as CF2-Net, Estan, FS-UNet and SMU-Net in the same dataset. The subjective visualization of the segmentation result of the proposed method is closest to the gold standards provided by experts, verifying that the proposed mehtod can segment the breast nodule area more accurately.  
摘要:  
针对超声成像分辨率低且存在噪声、结节形态与纹理复杂多变导致超声乳腺结节精确分割较为困难的问题,提出一种融合多尺度特征提取和跨空间特征融合的超声乳腺结节端到端自动分割方法。首先,设计一种多尺度特征提取与融合(MFEF)模块,通过融合4条具有不同感受野的卷积路径使网络具有多尺度的特征提取能力。其次,为对高级语义信息进行多尺度观察和信息筛选,在瓶颈层采用尺度感知与特征聚合(SFA)模块,以增强编码阶段的深层特征提取能力。此外,设计跨空间残差融合(CRF)模块,并将它应用在编、解码器间的跳跃连接上。该模块一方面对不同编码层进行跨空间信息融合,实现不同编码层间的信息互补;另一方面进一步提取编码层信息特征,缓解编解码对等层之间的语义差异,从而更好地补偿解码阶段的信息损失。在公开的超声乳腺结节数据集上的实验结果显示,所提方法的DICE系数可达0.888,同主流的深度学习分割模型UNet、AttUNet、ResUNet++、SKUNet相比,提高了0.033~0.094,对比相同数据集中的改进模型如CF2-Net、ESTAN、FS-UNet、SMU-Net,提高了0.001~0.068。所提方法分割结果图的主观视觉效果与专家给出的金标准最接近,能更加准确地分割出乳腺结节区域。  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1001-9081  
入藏号:  
CSCD:7625756  
Indexed Date:  
2024-03-09  
Institution Address:  
大连大学信息工程学院, 大连, 辽宁 116622, 中国  
Institution Address:  
大连市第五人民医院超声科, 大连, 辽宁 116021, 中国  
Institution Address:  
大连医科大学附属第一医院超声科, 大连, 辽宁 116011, 中国  
Institution Address:  
School of Information Engineering,Dalian University, Dalian, Liaoning 116622, China  
Institution Address:  
Department of Ultrasound,The Fifth People's Hospital of Dalian, Dalian, Liaoning 116021, China  
Institution Address:  
Department of Ultrasound,The First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning 116011, China  
Institution:  
大连大学信息工程学院  
Institution:  
大连市第五人民医院超声科  
Institution:  
大连医科大学附属第一医院超声科  
Institution:  
School of Information Engineering,Dalian University  
Institution:  
Department of Ultrasound,The Fifth People's Hospital of Dalian  
Institution:  
Department of Ultrasound,The First Affiliated Hospital of Dalian Medical University

### 321. YOLO-Faster: An efficient remote sensing object detection method based

摘要: As a pivotal task within computer vision, object detection finds application across a diverse spectrum of industrial scenarios. The advent of deep learning technologies has significantly elevated the accuracy of object detectors designed for general-purpose applications. Nevertheless, in contrast to conventional terrestrial environments, remote sensing object detection scenarios pose formidable challenges, including intricate and diverse backgrounds, fluctuating object scales, and pronounced interference from background noise, rendering remote sensing object detection an enduringly demanding task. In addition, despite the superior detection performance of deep learning-based object detection networks compared to traditional counterparts, their substantial parameter and computational demands curtail their feasibility for deployment on mobile devices equipped with low-power processors. In response to the aforementioned challenges, this paper introduces an enhanced lightweight remote sensing object detection network, denoted as YOLO-Faster, built upon the foundation of YOLOv5. Firstly, the lightweight design and inference speed of the object detection network is augmented by incorporating the lightweight network as the foundational network within YOLOv5, satisfying the demand for real-time detection on mobile devices. Moreover, to tackle the issue of detecting objects of different scales in large and complex backgrounds, an adaptive multiscale feature fusion network is introduced, which dynamically adjusts the large receptive field to capture dependencies among objects of different scales, enabling better modeling of object detection scenarios in remote sensing scenes. At last, the robustness of the object detection network under background noise is enhanced through incorporating a decoupled detection head that separates the classification and regression processes of the detection network. The results obtained from the public remote sensing object detection dataset DOTA show that the proposed method has a mean average precision of 71.4% and a detection speed of 38 frames per second.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0036-8504  
eISSN:  
2047-7163  
入藏号:  
WOS:001327108700001  
PubMed ID:  
39360473  
Indexed Date:  
2024-10-11  
Institution Address:  
Hangzhou Zhiyuan Res Inst Co Ltd, R&D Dept 4, Hangzhou, Peoples R China  
Institution Address:  
Zhejiang Univ, Polytech Inst, Hangzhou, Peoples R China  
Institution Address:  
Harbin Engn Univ, Qingdao Innovat & Dev Ctr, Qingdao, Peoples R China  
Institution Address:  
Harbin Engn Univ, Natl Key Lab Autonomous Marine Vehicle Technol, Harbin, Peoples R China  
Institution:  
Hangzhou Zhiyuan Res Inst Co Ltd  
Institution:  
Institution:  
Institution:

### 322. A Boundary Guided Cross Fusion Approach for Remote Sensing Image

摘要: Remote sensing images have a variety of application prospects because of their rich information. Due to recent advances in deep learning methods, solid improvements have been made in the semantic segmentation of high-resolution remote sensing images. However, achieving precise segmentation of small and crowded objects remains a challenge. To tackle this challenging task, a Boundary Guided Cross Fusion module (BGCFM) is proposed. The Bidirectional Boundary Gate module (BBGM) is designed to provide reliable boundary information for BGCFM. Based on these two models, a remote sensing images real-time semantic segmentation network, boundary guided cross fusion network (BGCFNet), is designed. The effectiveness of the boundary-guided fusion method and the performance of BGCFNet were verified on the GID-5 dataset without pretraining. The application of the boundary-guided fusion method on the SOTA dual-branch real-time semantic segmentation network improves segmentation accuracy. BGCFNet achieves the best performance with a Mean Intersection over Union (mIoU) of 88.82%. Its inference speed is about 1.5 times that of other networks in the experiment, achieving an excellent balance between accuracy and speed.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001166542700001  
Indexed Date:  
2024-03-06  
Institution Address:  
Changsha Univ Sci & Technol, Sch Comp & Commun, Changsha 410114, Peoples R China  
Institution:

### 323. AG-YOLO: Attention-guided network for real-time object detection

摘要: Existing neural network models directly add attention mechanisms to the network as a plug-and-play component to capture long-range dependencies and reconstruct feature maps. However, most methods do not fully tap the potential of attention in dealing with multi-scale problems. In this paper, an attention-guided YOLOv4 network (AG-YOLO) is proposed to address the multi-scale issue in object detection. We propose and apply multi-scale feature extraction to later stages of the backbone, which can not only enrich the feature hierarchy with low computational overhead, but also model the intra-scale and inter-scale correlation simultaneously to avoid missing key information. To reduce the redundant use of information flow, we propose a lightweight attention-guided feature pyramid network, which provides an efficient multi-level aggregation strategy based on multi-scale channel attention. In addition, a global context pathway is designed to reduce the dilution of high-level semantic information caused by information transmission. Compared with the baseline, AG-YOLO increased the mAP\_0.5 by 1.67%, while the number of parameters and GFLOPs merely increased by 0.33M and 0.18, respectively. Meanwhile, the detection accuracy of small object categories has been improved.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001058323600007  
Indexed Date:  
2023-09-19  
Institution Address:  
Southeast Univ, Sch Instrument Sci & Engn, Nanjing, Peoples R China  
Institution Address:  
Duke Kunshan Univ, Div Nat & Appl Sci, Suzhou, Peoples R China  
Institution:  
Institution:

### 324. TCSD: Triple Complementary Streams Detector for Comprehensive Deepfake

摘要: Advancements in computer vision and deep learning have made it difficult to distinguish deepfake visual media. While existing detection frameworks have achieved significant performance on challenging deepfake datasets, these approaches consider only a single perspective. More importantly, in urban scenes, neither complex scenarios can be covered by a single view nor can the correlation between multiple datasets of information be well utilized. In this article, to mine the new view for deepfake detection and utilize the correlation of multi-view information contained in images, we propose a novel triple complementary streams detector (TCSD). First, a novel depth estimator is designed to extract depth information (DI), which has not been used in previous methods. Then, to supplement depth information for obtaining comprehensive forgery clues, we consider the incoherence between image foreground and background information (FBI) and the inconsistency between local and global information (LGI). In addition, we designed an attention-based multi-scale feature extraction (MsFE) module to extract more complementary features from DI, FBI, and LGI. Finally, two attention-based feature fusion modules are proposed to adaptively fuse information. Extensive experiment results show that the proposed approach achieves state-of-the-art performance on detecting deepfakes.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
13  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
13  
ISSN:  
1551-6857  
eISSN:  
1551-6865  
入藏号:  
WOS:001035785200036  
Indexed Date:  
2023-08-24  
Institution Address:  
Beijing Jiaotong Univ, Inst Informat Sci, Beijing 100044, Peoples R China  
Institution Address:  
Beijing Key Lab Adv Informat Sci & Network Techno, Beijing 100044, Peoples R China  
Institution Address:  
Baidu Res, Inst Deep Learning, Beijing 100080, Peoples R China  
Institution Address:  
Baidu Res, Natl Engn Lab Deep Learning Technol & Applic, Beijing 100080, Peoples R China  
Institution Address:  
Univ Ubiquitous Co, IUNIUBI Res, Hangzhou 310000, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Univ Ubiquitous Co

### 325. Intelligent Identification and Prediction Mineral Resources Deposit

摘要: In recent years, the intelligent identification and prediction of ore deposits based on deep learning algorithm and image processing technology has gradually become one of the main research frontiers in the field of geological and metallogenic prediction. However, this method also has many problems that need to be solved. For example: (1) There are very few trainable image samples containing mineral point labels; (2) the geological image features are small and irregular, and the image similarity is high; (3) it is difficult to calculate the influence of different geological prospecting factors on ore mineralization. Based on this, this paper constructs a deep learning network model multiscale feature attention framework (MFAF) based on geoimage data. The results show that the MFCA-Net module in the MFAF model can solve the problem of scarce mine label images to a certain extent. In addition, the channel attention mechanism SE-Net module can quantify the difference in influence of different source factors on mineralization. The prediction map is obtained by applying the MFAF model in the study of deposit identification and prediction in the research area of the southern section of the Qin-hang metallogenic belt. The experimental results show that the areas numbered 5, 9, 16, 28, 34, 41, 50, 72, 74, 75, 80, 97, 101, 124, and 130 have great metallogenic potential and this method would be a promising tool for metallogenic prediction. A large number of experimental results show that this method has obvious advantages over other state-of-the-art methods in the prediction of prospecting target areas, and the prediction effect in the samples with mines is greatly improved. The multi-scale feature fusion and attention mechanism MFAF in this paper can provide a new way of thinking for geologists in mineral exploration. The research of this paper also provides resource guarantees and technical support for the sustainable exploitation of mineral resources and the sustainable growth of society and economy.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
eISSN:  
2071-1050  
入藏号:  
WOS:001028050800001  
Indexed Date:  
2023-07-28  
Institution Address:  
Wuyi Univ, Fac Intelligent Mfg, Jiangmen 529020, Peoples R China  
Institution Address:  
Chinese Acad Geol Sci, Inst Mineral Resources, Beijing 100037, Peoples R China  
Institution:  
Institution:

### 326. Video segmentation of industrial smoke based on dynamic fully

摘要: Accurately segmenting industrial smoke in videos plays a crucial role in assessing pollution levels based on smoke image evaluation. However, existing fully convolutional networks (FCNs) face challenges in precisely segmenting the edges of industrial smoke and exhibit low extraction and segmentation accuracy for small target smoke. To address this issue, we propose a video segmentation method specifically designed for industrial smoke. This method utilizes the dynamic FCN-Gaussian mixture model (GMM) along with a multi-scale fusion module and an attention module. The FCN-GMM effectively extracts dynamic feature information from spatiotemporal data, capturing motion in video or image sequences while preserving spatial details. The key innovation of FCN-GMM lies in integrating dynamic and static networks through a neural network, enabling the capture of features in both the temporal and spatial domains. Our approach begins by constructing a dynamic feature extraction network that captures spatial and temporal feature information separately during the training process, thereby enhancing the extraction of smoke edges. Additionally, we introduce a mechanism for multi-scale feature fusion and an attention module to effectively extract information related to small target smoke. Our experimental results demonstrate that our network accurately segments significant target smoke compared with FCNs. Furthermore, the network prioritizes the consideration of smoke edge information and improves the extraction of small target smoke, thereby enhancing the overall accuracy of smoke image segmentation with an increase of up to 10% in the intersection over union index. (C) 2023 SPIE and IS&T  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:001060595700045  
Indexed Date:  
2023-09-22  
Institution Address:  
Kunming Univ Sci & Technol, Fac Informat Engn & Automat, Kunming, Yunnan, Peoples R China  
Institution Address:  
Kunming Univ Sci & Technol, Yunnan Key Lab Artificial Intelligence, Kunming, Yunnan, Peoples R China  
Institution Address:  
Yunnan Kungang Elect & Inforrmat Sci Ltd, Kunming, Yunnan, Peoples R China  
Institution:  
Institution:  
Institution:  
Yunnan Kungang Elect & Inforrmat Sci Ltd

### 327. B-ultrasound guided venipuncture vascular recognition system based on

摘要: Venipuncture is one of the common operations used by doctors in the outpatient blood drawing room. The success rate of puncture is not only related to the pain degree of the patient, but also affects the test results. In this study, we propose an ultrasound guided venipuncture vascular recognition system based on deep learning. First, kmeans++ clustering is performed for the vascular regions in the different B-mode ultrasound images to facilitate estimation in subsequent work. Second, a lightweight vascular ultrasound network (UV-YOLOv7) is designed, specifically, based on YOLOv7-tiny, a multi-scale feature fusion module (MFFM) is designed to better fuse the high-level semantic features and low-level detail features, and the speed and accuracy of model detection are enhanced by lightweighting the model structure and replacing the EIoU loss function. Finally, a Dynamic Neighborhood-Density Based Spatial Clustering of Applications with Noise (DN-DBSCAN) algorithm is proposed, which can cluster a series of local vascular regions using the localization results and confidence properties of the network output to remove the misdetected regions. In the experiment, We selected 303 artifact-free and 264 heavily artifacted vascular ultrasound images for offline expansion and trained on the experimental platform, The results show that the proposed method performed best with an mAP of 86.2% and an inference time of 0.6 ms. At the end of the experiment, more robust vascular localization results were obtained by DN-DBSCAN clustering.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:001094492100001  
Indexed Date:  
2023-11-16  
Institution Address:  
Univ Shanghai Sci & Technol, Business Sch, Shanghai, Peoples R China  
Institution Address:  
Puncture Intelligent Med Technol Co, Shanghai, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Sch Mech Engn, Shanghai, Peoples R China  
Institution:  
Institution:  
Puncture Intelligent Med Technol Co  
Institution:

### 328. Advancing high-resolution remote sensing: a compact and powerful

摘要: Deep learning (DL)-based approaches are notable for their ability to establish feature associations without relying on physical constraints, unlike traditional strategies that are complex and dependent on expert experience. However, three main challenges hinder the versatility of semantic segmentation models. First, the targets in these images are dense and exist at varying spatial scales, which imposes higher demands on the model for accurate segmentation across scales. Second, the segmentation of small targets in the images is often overlooked, leading to a compromise between fine segmentation and model efficiency. Lastly, the data-intensive nature of remote sensing images and the resource-intensive operations of large-scale networks impose significant communication and computation burdens on edge devices, which may not have sufficient resources to handle them effectively. To address these challenges, this paper proposes a lightweight semantic segmentation method for remote sensing images to achieve high-precision segmentation for multi-scale targets while maintaining low computational complexity. The main components include: (1) embedding the inverted residual block structure to minimize the number of model parameters and computational costs; (2) introducing the parallel irregular space pyramid pooling module to efficiently aggregate multi-scale contextual information for fine-grained recognition of small targets; and (3) embedding transfer learning into the encoder-decoder structure to speed up the convergence rate and improve multi-scale feature fusion capability, thereby reducing semantic information loss. The proposed lightweight method has been extensively tested on real-world high-resolution remote sensing datasets. It achieved PA, MPA, MIoU, and FWIoU scores of 87.90%, 75.76%, 66.29%, and 78.81% on the Vaihingen dataset; 87.03%, 85.31%, 74.85%, and 77.54% on the Potsdam dataset; and 95.37%, 83.33%, 75.70%, and 91.31% on the Aeroscapes dataset. Compared to other popular semantic segmentation models, the proposed method achieved the highest values in all four evaluation indicators, demonstrating its effectiveness and superiority.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:001308392000001  
Indexed Date:  
2024-09-14  
Institution Address:  
Changchun Univ Sci & Technol, Sch Comp Sci & Technol, Changchun 130012, Peoples R China  
Institution Address:  
Changchun Inst Technol, Sch Comp Technol & Engn, Changchun, Peoples R China  
Institution Address:  
Changchun Inst Technol, Jilin Key Lab Changbai Hist Culture & VR Technol R, Changchun, Peoples R China  
Institution:  
Institution:  
Institution:

### 329. A multi-scale global attention network for blood vessel segmentation

摘要: Accurate segmentation of retinal fundus vessel images is vital to clinical diagnosis. Due to the intricate vascular morphology, high noise and low contrast of fundus vessel images, retinal fundus vessel segmentation is still a challenging task, especially for thin vessel segmentation. In recent years, on account of strong context feature extraction ability of deep learning, it has shown a remarkable performance in the automatic segmentation of retinal fundus vessels. However, it still exhibits certain limitations, such as information loss on micro objects or details, inadequate treatment of local features, etc. Faced with these challenging factors, we present a new multi-scale global attention network (MGA-Net). To realize effective feature representation, a dense attention U-Net network is proposed. Meanwhile, we design a global context attention (GCA) block to realize multi-scale feature fusion, allowing the global features from the deep network layers to flow to the shallow network layers. Further, aimed at retinal fundus vessel segmentation task again the class imbalance issue, the AG block is also introduced. Related experiments are conducted on CHASE\_DB1, DRIVE and STARE datasets to show the effectiveness of proposed segmentation model. The experimental results demonstrate the robustness of the proposed method with Ft exceeding 82% on all three datasets and effectively improve the segmentation performance of thin vessels. The source code of proposed MGA-Net is available at https://github.com/gegao310/workspace.git.  
Web of Science 核心合集中的 "被引频次":  
11  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
0263-2241  
eISSN:  
1873-412X  
入藏号:  
WOS:001088754000001  
Indexed Date:  
2023-11-07  
Institution Address:  
Zhengzhou Univ, Sch Elect & Informat Engn, Zhengzhou 450001, Henan, Peoples R China  
Institution Address:  
Robot Percept & Control Engn Lab, Zhengzhou 450001, Henan, Peoples R China  
Institution Address:  
Zhengzhou Univ Light Ind, Coll Comp & Commun Engn, Zhengzhou 450001, Peoples R China  
Institution:  
Institution:  
Robot Percept & Control Engn Lab  
Institution:

### 330. MFFENet and ADANet: a robust deep transfer learning method and its

摘要: Automatic recognition and segmentation methods have become an essential requirement in identifying large-scale earthquake-induced landslides. This used to be conducted through pixel-based or object-oriented methods. However, these methods fail to develop an accurate, rapid, and cross-scene solution for earthquake-induced landslide recognition because of the massive amount of remote sensing data and variations in different earthquake scenarios. To fill this research gap, this paper proposes a robust deep transfer learning scheme for high precision and fast recognition of regional landslides. First, a Multi-scale Feature Fusion regime with an Encoder-decoder Network (MFFENet) is proposed to extract and fuse the multi-scale features of objects in remote sensing images, in which a novel and practical Adaptive Triangle Fork (ATF) Module is designed to integrate the useful features across different scales effectively. Second, an Adversarial Domain Adaptation Network (ADANet) is developed to perform different seismic landslide recognition tasks, and a multi-level output space adaptation scheme is proposed to enhance the adaptability of the segmentation model. Experimental results on standard remote sensing datasets demonstrate the effectiveness of MFFENet and ADANet. Finally, a comprehensive and general scheme is proposed for earthquake-induced landslide recognition, which integrates image features extracted from MFFENet and ADANet with the side information including landslide geologic features, bi-temporal changing features, and spatial analysis. The proposed scheme is applied in two earthquake-induced landslides in Jiuzhaigou (China) and Hokkaido (Japan), using available pre- and post-earthquake remote sensing images. These experiments show that the proposed scheme presents a state-of-the-art performance in regional landslide identification and performs stably and robustly in different seismic landslide recognition tasks. Our proposed framework demonstrates a competitive performance for high-precision, high-efficiency, and cross-scene recognition of earthquake disasters, which may serve as a new starting point for the application of deep learning and transfer learning methods in earthquake-induced landslide recognition.  
在 BIOSIS Citation Index 中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
30  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
5  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
34  
ISSN:  
1612-510X  
eISSN:  
1612-5118  
入藏号:  
WOS:000776423800003  
Indexed Date:  
2022-04-14  
Institution Address:  
Chinese Acad Sci, Inst Mt Hazards & Environm, Key Lab Mt Hazards & Surface Proc, Chengdu 610041, Peoples R China  
Institution Address:  
Chinese Acad Sci, CAS Ctr Excellence Tibetan Plateau Earth Sci, Beijing 100101, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution Address:  
Westlake Univ, Hangzhou 310024, Peoples R China  
Institution Address:  
Chengdu Univ Technol, State Key Lab Geohazard Prevent & Geoenvironm Pro, Chengdu 610059, Peoples R China  
Institution Address:  
Sichuan Geomat Ctr, Sichuan Engn Res Ctr Emergency Mapping & Disaster, Chengdu 610041, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Sichuan Geomat Ctr

### 331. Learning Template-Constraint Real-Time Siamese Tracker for Drone AI

摘要: Significant progress has been made in object tracking tasks thanks to the application of deep learning. However, current deep neural network-based object tracking methods often rely on stacking sub-modules and introducing complex structures to improve tracking accuracy. Unfortunately, these approaches are inefficient and limit the feasibility of deploying efficient trackers on drone AI devices. To address these challenges, this paper introduces ConcatTrk, a high-speed object tracking method designed specifically for drone AI devices. ConcatTrk utilizes a lightweight network architecture, enabling real-time tracking on edge devices. Specifically, the proposed method primarily uses the concatenation operation to construct its core tracking steps, including multi-scale feature fusion, intra-frame feature matching, and dynamic template updating, which aim to reduce the computational overhead of the tracker. To ensure tracking performance in UAV tracking scenarios, ConcatTrk implements a learnable feature matching operator along with a simple and efficient template constraint branch, which enables accurate tracking by discriminatively matching features and incorporating periodic template updates. Results of comprehensive experiments on popular benchmarks, including UAV123, OTB100, and LaSOT, show that ConcatTrk has achieved promising accuracy and attained a tracking speed of 41 FPS on an edge AI device, Nvidia AGX Xavier. ConcatTrk runs 8x faster than the SOTA tracker TransT while using 4.9x fewer FLOPs. Real-world tests on the drone platform have strongly validated its practicability, including real-time tracking speed, reliable accuracy, and low power consumption.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2504-446X  
入藏号:  
WOS:001077125400001  
Indexed Date:  
2023-10-18  
Institution Address:  
Univ Elect Sci & Technol China, Sch Informat & Software Engn, 4,Sect 2,Jianshebei Rd, Chengdu 610054, Peoples R China  
Institution:

### 332. BMSFormer: An efficient deep learning model for online state-of-health

摘要: The efficient and accurate state-of-health (SOH) estimation is crucial for reducing risks and ensuring effective application in battery management systems (BMS) of resource-limited devices. However, many recent state-ofthe-art SOH estimation approaches rely on resource-consuming structures to obtain good performance. In this paper, an efficient deep learning model for SOH estimation, namely BMSFormer, is constructed. BMSFormer mainly integrates a Local-Global Fusion Attention module to capture both long-term and short-term dependencies while reducing computational complexity compared to traditional Softmax-based attention. Additionally, two kinds of depthwise separable convolution are embedded to fuse multi-scale and multi-channel features, enhancing feature diversity with fewer parameters than standard convolution. Three widely used battery datasets, each with different chemistries and operating conditions, are employed to evaluate the performance of BMSFormer. The experiments results illustrate that the proposed model achieves higher accuracy, lower computational consumption, and stabler performance across various hyperparameter combinations compared to alternative models.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0360-5442  
eISSN:  
1873-6785  
入藏号:  
WOS:001373539800001  
Indexed Date:  
2024-12-16  
Institution Address:  
Harbin Inst Technol, Dept Mech Engn, Weihai 264209, Shandong, Peoples R China  
Institution Address:  
Harbin Inst Technol, Sch Mechatron Engn, Harbin 150000, Heilongjiang, Peoples R China  
Institution Address:  
Harbin Inst Technol, Weihai Key Lab Intelligent Operat & Maintenance, Weihai 264209, Shandong, Peoples R China  
Institution Address:  
Harbin Inst Technol, Sch Automot Engn, Weihai 264209, Shandong, Peoples R China  
Institution Address:  
Civil Aviat Univ China, Coll Aeronaut Engn, Tianjin 300300, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 333. Synthetic aperture optical image restoration based on multi-scale

摘要: With the wide applications of high-resolution imaging technology in topographic mapping, astronomical observation, and military reconnaissance and other fields, the requirements for imaging resolution of optical system are becoming higher and higher. According to the diffraction limit and Rayleigh criterion, the imaging resolution of the optical system is proportional to the size of the aperture of the system, but affected by the material and the processing of the optical component: the single aperture of the optical system cannot be infinitely enlarged. Therefore the synthetic aperture technology is proposed to replace the single large aperture optical system. Owing to the effect of sub-aperture arrangement and light scattering, the imaging of synthetic aperture optical system will be degraded because of insufficient light area and phase distortion. The traditional imaging restoration algorithm of synthetic aperture optical system is sensitive to noise, overly relies on degraded model, requires a lot of manually designed models, and has poor adaptability. To solve this problem, a multiscale feature enhancement method of restoring the synthetic aperture optical image is proposed in this work. U-Net is used to obtain multi-scale feature, and self-attention in mixed domain is used to improve the ability of of the network to extract the features in space and channel. Multi-scale feature fusion module and feature enhancement module are constructed to fuse the information between features on different scales. The information interaction mode of the codec layer is optimized, the attention of the whole network to the real structure of the original image is enhanced, and the artifact interference caused by ringing is avoided in the process of restoration. The final experimental results are 1.51%, 4.42% and 5.22% higher than those from the advanced deep learning algorithms in the evaluation indexes of peak signal-to-noise ratio, structural similarity and perceived similarity, respectively. In addition, the method presented in this work has a good restoration effect on the degraded images to different degrees of synthetic aperture, and can effectively restore the degraded images and the images with abnormal light, so as to solve the problem of imaging degradation of synthetic aperture optical system. The feasibility of deep learning method in synthetic aperture optical image restoration is proved.  
[GRAPHICS]  
.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1000-3290  
入藏号:  
WOS:001195279300014  
Indexed Date:  
2024-04-17  
Institution Address:  
Wuxi Univ, Jiangsu Prov Engn Res Ctr Integrated Circuit Reli, Wuxi 214105, Jiangsu, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Elect & Informat Engn, Nanjing 210044, Peoples R China  
Institution:  
Institution:

### 334. A Flexible Monitoring Framework via Dynamic-Multilayer Graph Convolution

摘要: Due to the continuous technological innovation in industrial processes, many deep learning-based methods have shown powerful capability in handling equipment status monitoring, but most of them ignore the temporal features and the dynamic changes of the diverse spatial structure of the raw data. Meanwhile, these methods usually focus on handling a single downstream task but rarely consider different tasks simultaneously. To solve these issues, this article proposes a more flexible monitoring framework based on a dynamic-multilayer graph convolution network (GCN), which can be adapted to different downstream tasks simultaneously by agile combinations of the modules according to the different industrial scenarios. First, the time series of fault samples are segmented and constructed into a feature matrix to extract the temporal information by the temporal module. At the same time, to fully characterize the dynamic change of different spatial structures among samples, the dynamic graph of each moment is expanded into a multilayer graph through various composition indexes. Then, the spatial information is extracted by the intralayer and interlayer convolution operations in the spatial module, and the fused features are applied to different downstream tasks. Finally, the experiments are performed in two different downstream tasks, namely, three datasets for fault diagnosis and one dataset for remaining useful life (RUL) prediction. The results of both diagnosis and prediction are better than those of the comparison algorithms, so the effectiveness and flexibility of the proposed framework are proven.  
Web of Science 核心合集中的 "被引频次":  
26  
在中国科学引文数据库中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
27  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001029190800030  
Indexed Date:  
2024-08-30  
Institution Address:  
Anhui Univ, Sch Math Sci, Key Lab Intelligent Comp & Signal Proc, Minist Educ, Hefei 230601, Peoples R China  
Institution Address:  
Anhui Univ, Inst Phys Sci & Informat Technol, Key Lab Intelligent Comp & Signal Proc, Minist Educ, Hefei 230601, Peoples R China  
Institution:  
Institution:

### 335. Analysis of children's sub-health treatment effect based on multi-scale

摘要: Sub-health state is a state of health and low quality between disease and health. The theoretical basis of children's sub-health is to start from the whole. The common clinical sub-health conditions cannot be explained by modern detection methods, and it can be screened and analyzed with the help of big data in medical informatization. The combination of "Internet + " and the health care model is an innovation in the construction of medical informatization. It can provide many considerate services to the masses in time and alleviate the anxiety of illness. Therefore, it is very necessary to carry out the efficacy evaluation of children's sub-health from the perspective of medical information. Therefore, this paper completes the following work with the help of AI neural network: (1) This paper proposes an improved AlexNet network evaluation method based on attention mechanism. In this study, attention mechanism is added to the original AlexNet model to weight each channel of the feature layer. At the same time, we improve the large convolution kernel of the previous layers of the original AlexNet network and use batch normalization instead of the local response normalization (LRN) layer in the original model. (2) This paper proposes an evaluation method based on improved residual network, which improves the original residual block of the residual model and widens the residual block. The residual block can effectively reduce the amount of network parameters and improve the efficiency of network training. (3) This paper proposes an evaluation method of multi-scale feature fusion (MSFF). The features extracted from the improved AlexNet and residual network are fused and then evaluated. At this time, the training time is greatly shortened, and the accuracy is higher than that of single model.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0941-0643  
eISSN:  
1433-3058  
入藏号:  
WOS:001089027100001  
Indexed Date:  
2023-11-06  
Institution Address:  
Henan Univ Chinese Med, Henan Prov Hosp TCM, Dept Pediat, Zhengzhou 450000, Henan, Peoples R China  
Institution:

### 336. Wave-Net: A lightweight deep network for retinal vessel segmentation

摘要: Accurate segmentation of retinal vessels from fundus images is fundamental for the diagnosis of numerous diseases of eye, and an automated vessel segmentation method can effectively help clinicians to make accurate diagnosis for the patients and provide the appropriate treatment schemes. It is important to note that both thick and thin vessels play the key role for disease judgements. Because of complex factors, the precise segmentation of thin vessels is still a great challenge, such as the presence of various lesions, image noise, complex backgrounds and poor contrast in the fundus images. Recently, because of the advantage of context feature representation learning capabilities, deep learning has shown a remarkable segmentation performance on retinal vessels. However, it still has some shortcomings on high-precision retinal vessel extraction due to some factors, such as semantic information loss caused by pooling operations, limited receptive field, etc. To address these problems, this paper proposes a new lightweight segmentation network for precise retinal vessel segmentation, which is called as Wave-Net model on account of the whole shape. To alleviate the influence of semantic information loss problem to thin vessels, to acquire more contexts about micro structures and details, a detail enhancement and denoising block (DED) is proposed to improve the segmentation precision on thin vessels, which replaces the simple skip connections of original U-Net. On the other hand, it could well alleviate the influence of the semantic gap problem. Further, faced with limited receptive field, for multi-scale vessel detection, a multi-scale feature fusion block (MFF) is proposed to fuse cross-scale contexts to achieve higher segmentation accuracy and realize effective processing of local feature maps. Experiments indicate that proposed Wave-Net achieves an excellent performance on retinal vessel segmentation while maintaining a lightweight network design compared to other advanced segmentation methods, and it also has shown a better segmentation ability to thin vessels.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
45  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
8  
被引频次合计:  
47  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:000921201200001  
PubMed ID:  
36463794  
Indexed Date:  
2023-03-09  
Institution Address:  
Zhengzhou Univ, Sch Elect & Informat Engn, Zhengzhou 450001, Peoples R China  
Institution Address:  
Robot Percept & Control Engn Lab, Kaifeng 450001, Henan, Peoples R China  
Institution Address:  
Edinburgh Napier Univ, Built Environm, Edinburgh EH10 5DT, Scotland  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Management & Control Complex Syst, Beijing 100190, Peoples R China  
Institution:  
Institution:  
Robot Percept & Control Engn Lab  
Institution:  
Institution:

### 337. SC-Net: A new strip convolutional network model for rice seedling and

摘要: Weeds are among the major factors that could harm the yield and quality of rice. Accurately recognizing and localizing crops and weeds are essential for realizing automated weed management in precision agriculture. Semantic segmentation techniques based on deep learning have the capability to automatically discern various types of objects. However, effectively extracting image features to distinguish between rice seedlings and weeds, which often exhibit similar texture characteristics and size disparities, remains a challenging issue in the field. In view of this, a new strip convolutional network model named SC-Net is proposed in this paper, where UNet is used as the backbone network. Based on the idea of multi-scale feature fusion, the parallel multilevel convolution block (PMCB) and strip multilevel convolution block (SMCB) are constructed to design the encoder and decoder of the segmentation network, enabling the extraction of the salient features of seedlings and weeds. Specifically, the SMCB is composed of multi-scale strip convolutions, which effectively widens the receptive field of the convolution layer while minimizing computational costs, and incorporates a long and narrow shape enhancement network to identify the characteristics of slender rice seedling leaves. To adaptively fuse different level features, the attention feature fusion module (AFF) is designed to establish a long skip connection between the encoder and decoder of the network. This module aggregates global and local context information from low-level and high-level features through global spatial pooling and dot product convolution. Moreover, the strip pooling attention module (SPAM) is introduced between the encoder and decoder stages to enhance the network's perception of the precise positional information of the target area, thus further optimizing the segmentation results. The experimental results show that SC-Net achieved MIOU scores of 87.48 % and 89.00 % on the selfbuilt rice seedling and public agricultural datasets. Compared with several state-of-the-art models, the proposed model achieves better segmentation performance, thus contributing to providing a promising support for the development of intelligent weeding in the field.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在 BIOSIS Citation Index 中的被引频次:  
3  
被引频次合计:  
6  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:001224721600001  
Indexed Date:  
2024-05-23  
Institution Address:  
Anhui Agr Univ, Sch Engn, Hefei 230036, Peoples R China  
Institution Address:  
Univ Leicester, Sch Comp & Math Sci, Leicester, England  
Institution Address:  
Anhui Acad Agr Sci, Rice Res Inst, Hefei 230031, Peoples R China  
Institution:  
Institution:  
Institution:

### 338. IR-MSDNet: Infrared and Visible Image Fusion Based On Infrared Features

摘要: Infrared (IR) and visible images are heterogeneous data, and their fusion is one of the important research contents in the remote sensing field. In the last decade, deep networks have been widely used in image fusion due to their ability to preserve high-level semantic information. However, due to the lower resolution of IR images, deep learning-based methods may not be able to retain the salient features of IR images. In this article, a novel IR and visible image fusion based on IR Features & Multiscale Dense Network (IR-MSDNet) is proposed to preserve the content and key target features from both visible and IR images in the fused image. It comprises an encoder, a multiscale decoder, a traditional processing unit, and a fused unit, and can capture incredibly rich background details in visible images and prominent target details in IR features. When the dense and multiscale features are fused, the background details are obtained by utilizing attention strategy, and then combined with complimentary edge features. While IR features are extracted by traditional quadtree decomposition and Bezier interpolation, and further intensified by refinement. Finally, both the decoded multiscale features and IR features are used to reconstruct the final fused image. Experimental evaluation with other state-of-the-art fusion methods validates the superiority of our proposed IR-MSDNet in both subjective and objective evaluation metrics. Additional objective evaluation conducted on the object detection (OD) task further verifies that the proposed IR-MSDNet has greatly enhanced the details in the fused images, which bring the best OD results.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
21  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
22  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000637532600010  
Indexed Date:  
2021-05-04  
Institution Address:  
Shanghai Jiao Tong Univ, Dept Automat, Shanghai 200240, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Key Lab Syst Control & Informat Proc, Minist Educ, Shanghai 200240, Peoples R China  
Institution Address:  
Sch Geog Sci, Minist Educ, Key Lab Geog Informat Sci, Shanghai 200241, Peoples R China  
Institution Address:  
East China Normal Univ, Shanghai Key Lab Multidimens Informat, Shanghai 200062, Peoples R China  
Institution:  
Institution:  
Institution:  
Sch Geog Sci  
Institution:

### 339. High-resolution feature pyramid attention network for high spatial

摘要: Land-cover classification based on remote sensing technology has been adopted for decision-making concerning agricultural development, urban planning, and ecosystem protection in arid oasis zones. The semantic segmentation method based on deep learning, as a new paradigm, can effectively overcome the limitations of traditional pixel-based and object-based methods and obtain good classification results from high spatial resolution (HSR) remote sensing images. However, how to extract the exact category boundary and realize the high precision mapping is still a problem. This paper proposes a novel high-resolution feature pyramid attention network (HRFPANet) for land-cover classification. It effectively integrates the advantages of multi-scale feature extraction, attention mechanism, and feature fusion and alleviates boundary inconsistency, roughness, and category fragmentation associated with previous semantic segmentation models. The experimental results show that the mIoU score of HRFPANet is 79.5%, which is 11.5% and 2.6% higher than that of PSPNet and UPerNet, respectively. It proves the proposed model can be used for qualified land-cover mapping in arid oasis zones. Our source code is available at https://github.com/HPU-CPD/HRFPANet.git.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:001229887400001  
Indexed Date:  
2024-05-31  
Institution Address:  
Lanzhou Univ, Coll Earth & Environm Sci, Lanzhou, Peoples R China  
Institution Address:  
Mapping Inst Gansu Prov, Dept Remote Sensing & Geog Natl Condit Monitoring, Lanzhou, Peoples R China  
Institution:  
Institution:  
Mapping Inst Gansu Prov

### 340. Self-supervised few-shot medical image segmentation with multi-attention

摘要: Mainstream fully supervised deep learning segmentation models can achieve good results when trained on abundant labeled data, but the image segmentation in the medical field faces the challenges of high annotation cost and diverse segmentation targets, often lacking sufficient labeled data. The model proposed in this paper incorporates the idea of extracting labels from data through self-supervision, utilizing superpixels to represent image characteristics for image segmentation under conditions of small sample annotation. The introduction of multiple attention mechanisms allows the model to focus more on spatial features of the image. The position attention module and channel attention module aim to fuse multi-scale features within a single image, while the external attention module highlights the connections between different samples. Experiments were conducted on the CHAOS healthy abdominal organ dataset. In the extreme case of the 1-shot, DSC reached 0.76, which is about 3%higher than the baseline result. In addition, this paper explores the significance of few-shot learning by adjusting the number of N-way-K-shot tasks. Under the 7-shot setting, DSC achieves significant improvement, which is within an acceptable range of the segmentation effect based on full supervision based on deep learning.  
摘要:  
主流的基于全监督的深度学习分割模型在丰富的标记数据上训练时可以取得良好的效果,但医疗图像领域的图像分割存在标注成本高、分割目标种类多的问题,且往往缺少足够的标注数据。提出一个模型,通过融合自监督从数据中提取标签,利用超像素表征图像特性,进行小样本标注条件下的图像分割。引入多注意力机制使得模型更多关注图像的空间特征,位置注意模块和通道注意模块致力于单一图像内部的多尺度特征融合,而外部注意力模块显著突出了不同样本间的联系。在CHAOS健康腹部器官数据集上进行实验,1-shot极端情况下DSC达0.76,相较baseline分割结果提升3%左右。通过调整Nway- K-shot任务数来探讨小样本学习的意义,在7-shot设置下DSC有显著提升,与基于全监督的深度学习分割效果的差距在可接受范围内。  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1007-130X  
入藏号:  
CSCD:7683020  
Indexed Date:  
2024-06-07  
Institution Address:  
太原理工大学计算机科学与技术学院(大数据学院), 晋中, 山西 030600, 中国  
Institution Address:  
太原理工大学计算机科学与技术学院(大数据学院);;长治医学院, ;;山西省智能数据辅助诊疗工程研究中心, 晋中;;长治, 山西;;山西 030600;;046000, 中国  
Institution Address:  
College of Computer Science and Technology (College of Data Science),Taiyuan University of Technology, Jinzhong, Shanxi 030600, China  
Institution Address:  
College of Computer Science and Technology (College of Data Science),Taiyuan University of Technology;;Changzhi Medical College, ;;Shanxi Engineering Research Centre for Intelligent Data Assisted Treatment, Jinzhong;;Changzhi, ;; 030600;;046000  
Institution:  
太原理工大学计算机科学与技术学院(大数据学院)  
Institution:  
太原理工大学计算机科学与技术学院(大数据学院);;长治医学院  
Institution:  
College of Computer Science and Technology (College of Data Science),Taiyuan University of Technology  
Institution:  
College of Computer Science and Technology (College of Data Science),Taiyuan University of Technology;;Changzhi Medical College

### 341. Hybrid Network Based on Hierarchical Multipatch Feature Encoder for

摘要: With recent advances in artificial intelligence, deep learning networks with various structures are being applied to target detection. However, in the infrared small target detection (IRSTD) field, an appropriate network structure is still required to identify blurry and low-contrast targets accurately. Conventional U-net algorithms use skip connection and attention module because information loss occurs as the convolution layer becomes deeper. However, the problem of not being able to recognize the entire image due to the limitations of convolution is fatal in IRSTD. To overcome these limitations, we design an encoder that divides the image into multiple small pieces and stacks them hierarchically to extract features. Therefore, the proposed network has a hybrid encoder structure that combines a convolution-based multiscale encoder to extract local information and a hierarchical-based multipatch encoder to extract global information by running in parallel. To effectively fuse the hierarchical information obtained from each layer, the loss function, which plays an important role in learning, has also been changed to suit the hybrid encoder. The proposed algorithm, named hierarchical multipatch feature and attention multiscale feature fusion U-net (HMAMFU-net), can guarantee effective target detection performance with two datasets: NUDT-SIRST and NUAA-SIRST, and performance analysis is conducted with other state-of-the-art algorithms. PyTorch implementation is available at https://github.com/skylih87/HMAMFU-net.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001308225500008  
Indexed Date:  
2024-09-18  
Institution Address:  
Seoul Natl Univ, Automat & Syst Res Inst, Dept Aerosp Engn, Seoul 08826, South Korea  
Institution:

### 342. Learning Motion Constraint-Based Spatio-Temporal Networks for Infrared

摘要: Efficient infrared dim object detection has been challenged by low signal-to-noise ratios (SNRs). Traditional methods rely on the gradient difference and fixed-parameter model. These methods fail to adapt to sophisticated and variable situations in the real world. To tackle the issue, a deep learning method based on the spatio-temporal network is proposed in this paper. The model is established by the Convolutional Long Short-Term Memory cell (Conv-LSTM) and the 3D Convolution cell (3D-Conv). It is trained to learn the motion constraint of moving targets (spatio-temporal constraint module, called STM) and to fuse the multiscale local feature between the target and background (deep spatial features module, called DFM). In addition, a variable interval search module (state-aware module, called STAM) is added to the inference. The submodule decides to conduct a global search for images only if the target is lost due to fast motion, uncertain obstruction, and frame loss. Comprehensive experiments indicate that the proposed method achieves better performance over all baseline methods. On the mid-wave infrared datasets collected by the authors, the proposed method achieves a 95.87% detection rate. The SNR of the dataset is around 1-3 dB, and the background of the sequence includes sky, asphalt road, and buildings.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2076-3417  
入藏号:  
WOS:000887134900001  
Indexed Date:  
2022-12-05  
Institution Address:  
Chinese Acad Sci, Shanghai Inst Tech Phys, Key Lab Intelligent Infrared Percept, Shanghai 200083, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution Address:  
Shanghai Tech Univ, Sch Informat Sci & Technol, Shanghai 201210, Peoples R China  
Institution:  
Institution:  
Institution:

### 343. Local Patch Network With Global Attention for Infrared Small Target

摘要: Infrared small target detection plays an important role in the infrared search and tracking applications. In recent years, deep learning techniques have been introduced to this task and achieved noteworthy effects. Following general object segmentation methods, existing deep learning methods usually process the image from the global view. However, the locality of small targets and extreme class-imbalance between the target and background pixels are not well-considered by these deep learning methods, which causes the low-efficiency on training and high-dependence on numerous data. A local patch network (LPNet) with global attention is proposed in this article to detect small targets by jointly considering the global and local properties of infrared small target images. From the global view, a supervised attention module trained by the small target spread map is proposed to suppress most background pixels irrelevant with small target features. From the local view, local patches are split from global features and share the same convolution weights with each other in an LPNet. By leveraging both the global and local properties, the data-driven framework proposed in this article has the ability of fusing multiscale features for small target detection. Extensive experiments on synthetic and real datasets show that the proposed method achieves the state-of-the-art performance in comparison with both traditional and deep learning methods.  
Web of Science 核心合集中的 "被引频次":  
39  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
43  
ISSN:  
0018-9251  
eISSN:  
1557-9603  
入藏号:  
WOS:000866503000021  
Indexed Date:  
2022-10-26  
Institution Address:  
Chongqing Univ Posts & Telecommun, Sch Commun & Informat Engn, Chongqing, Peoples R China  
Institution Address:  
Chongqing Key Lab Signal & Informat Proc, Chongqing, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Sch Math & Stat, Xian 710049, Peoples R China  
Institution Address:  
Macau Univ Sci & Technol, Fac Informat Technol, Macau, Peoples R China  
Institution Address:  
Harbin Inst Technol, Sch Comp Sci & Technol, Harbin 150001, Peoples R China  
Institution Address:  
Macau Univ Sci & Technol, Macau Inst Syst Engn, Taipa 999078, Macau, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Chongqing Key Lab Signal & Informat Proc, Chongqing 400065, Peoples R China  
Institution:  
Institution:  
Chongqing Key Lab Signal & Informat Proc  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 344. Research on Bone Stick Text Recognition Method with Multi-Scale Feature

摘要: Bone sticks are composed of thin slices of animal bones created by ancient people, which mainly served the functions of fixing books, writing scripts, and divination. The bone stick script is an essential material for studying the history of Chinese Western Han script. Using a neural network for text recognition can quickly interpret ancient text, while extracting deeper semantic information, neural networks will also lose superficial image details. After multi-layer convolution and pooling of bone sticks, the continuous loss of superficial details affects classification accuracy. At the same time, the unbalanced distribution of bone stick quantity leads to a low recognition rate with small samples of bone sticks. Aiming to solve the above problems, a bone stick recognition method based on multi-scale features and focal loss function is proposed. Firstly, based on the residual network ResNet, the output features of the first layer and four Conv\_x layers are pooled globally to reduce the feature dimension of each channel, and the channel splicing method is used to add different depths of base information to the original high-level features, which improves the detail feature extraction ability of the model. Secondly, in view of the unbalanced distribution of the bone stick data, the original cross-entropy loss function is replaced by the focus loss function, which increases the penalty for classification errors and improves the recognition rate of classes with few training samples. Experimental results show that the recognition accuracy of the proposed method on the bone stick data set is up to 90.5%.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:000900456000001  
Indexed Date:  
2023-01-04  
Institution Address:  
Xian Univ Architecture & Technol, Sch Informat & Control Engn, Xian 710055, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Archaeol, Beijing 100101, Peoples R China  
Institution:  
Institution:

### 345. Automated Pavement Distress Detection Based on Convolutional Neural

摘要: Pavement distress detection is crucial in road health assessment and monitoring. However, there are still some challenges in extracting pavement distress based on deep learning: such as insufficient segmentation, extraction errors and discontinuities. In this paper, we propose DARNet, a network for pavement distress extraction. A Distress Aware Attention Module (DAAM) is proposed to solve the problem of discontinuity in distress extraction due to inaccurate recovery of distress pixels during upsampling. Based on the characteristics of distress morphology, a Refinement Extraction Module (REM) is designed to effectively capture horizontal and vertical pavement damage features by fusing high-level and low-level features, which improves the accuracy of the model in extracting details of pavement damage information. Finally, a Weighted Cross-Entropy Loss function (WCEL) is introduced to assign weights according to the distance of the pixel point to the boundary of the distress, which solves the problem that the traditional cross entropy function treats each pixel point equally. We also propose a set of pavement distress datasets LNTU\_RDD\_GS, and the experimental results show that DARNet can reach 82.68% mIoU and 90.13% F score in the datasets in this paper, 80.63% mIoU and 88.35% F score in the four public datasets.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001286669800001  
Indexed Date:  
2024-08-13  
Institution Address:  
Liaoning Tech Univ, Sch Geomat, Fuxin 123000, Peoples R China  
Institution Address:  
Liaoning Tech Univ, Collaborat Innovat Inst Geospatial Informat Serv, Fuxin 123000, Peoples R China  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & Re, Wuhan 430079, Peoples R China  
Institution:  
Institution:  
Institution:

### 346. Multi-Scale Attention-Enhanced Deep Learning Model for Ionogram

摘要: To increase the accuracy of ionogram automatic scaling, a deep learning model-multi-scale attention-enhanced UNet is proposed. Correspondingly, a multi-scale attention-enhanced (MSAE) sub-network is developed which involves a spatial attention nearest up-sampling module and several residual channel attention modules with multi-scale skip connections. They contribute to multi-scale feature fusion and augmentation of the learning ability for enhancing faint and elongated profile traces of ionograms. The MSAE sub-network input consists of multi-scale feature maps which could be optimally employed to make the network effectively utilize useful information from the encoders and decoders. Incidentally, a dual channel spatial attention block is embedded between the encoder and the decoder for deeper detail extraction. When the proposed model is applied to scale different electron density profiles of ionograms based on an open data set, the experimental results show the segmentation performance evaluation indexes: the precision and the recall rate can be improved by 6.9% and 26.1%, respectively, compared to automatic real-time ionogram scaling with true-height routine. Another set of indexes: the mean intersection over union and the F-score are superior to that of other several contrasted deep learning models, which can be improved by 3% and 1.6%, respectively, compared to the original UNet model.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0048-6604  
eISSN:  
1944-799X  
入藏号:  
WOS:000949549100001  
Indexed Date:  
2023-07-03  
Institution Address:  
Xiangtan Univ, Sch Automat & Elect Informat, Xiangtan, Peoples R China  
Institution:

### 347. Few-Shot PolSAR Ship Detection Based on Polarimetric Features Selection

摘要: Deep learning methods have been widely studied in the field of polarimetric synthetic aperture radar (PolSAR) ship detection over the past few years. However, the backscattering of manmade targets, including ships, is sensitive to the relative geometry between target orientation and radar line of sight, which makes the diversity of polarimetric and spatial features of ships. The diversity of scattering leads to a relative increase in the scarcity of PolSAR-labeled samples, which are difficult to obtain. To solve the abovementioned issue and extract the polarimetric and spatial features of PolSAR images better, this paper proposes a few-shot PolSAR ship detection method based on the combination of constructed polarimetric input data selection and improved contrastive self-supervised learning (CSSL) pre-training. Specifically, eight polarimetric feature extraction methods are adopted to construct deep learning network input data with polarimetric features. The backbone is pre-trained with un-labeled PolSAR input data through an improved CSSL method without negative samples, which enhances the representation capability by the multi-scale feature fusion module (MFFM) and implements a regularization strategy by the mix-up auxiliary pathway (MUAP). The pre-trained backbone is applied to the downstream ship detection network; only a few labeled samples are used for fine-tuning and the construction method of polarimetric input data with the best detection effect is studied. The comparison and ablation experiment results on the self-established PolSAR ship detection dataset verify the superiority of the proposed method, especially in the case of few-shot learning.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
eISSN:  
2072-4292  
入藏号:  
WOS:000969861900001  
Indexed Date:  
2023-05-05  
Institution Address:  
Beihang Univ, Sch Elect & Informat Engn, Beijing 100191, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100190, Peoples R China  
Institution Address:  
Chinese Acad Sci, Key Lab Technol Geospatial Informat Proc & Applica, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 101408, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 348. HD-Net: High-resolution decoupled network for building footprint

摘要: The extraction of building footprints, as a highly challenging task in remote sensing (RS) image -based geospatial object detection and recognition, holds significant importance. Due to the strong coupling in RS images between the body and boundary of buildings, the ability of most currently advanced deep learning models in building footprint extraction remains limited, inevitably meeting the extraction performance bottleneck. To this end, we propose a novel High -resolution Decoupled Network, HD -Net for short, for precious building footprint extraction in RS. HD -Net follows the well-known high -resolution network (HRNet) architecture, which can to a great extent alleviate the coupling issues between body and boundary using its multi -scale information interaction in parallel. More specifically, Our HD -Net innovatively designs the multiple stacked multi -scale feature fusion (MFF) modules, where the MFF module is performed by combining the deep supervision technique and a feature decoupling-recoupling (FDR) module. The FDR module adeptly untangles coupled features into two distinct elements: body and boundary, yielding feature maps enriched with semantic information. This configuration facilitates a step -wise refinement of building extraction and boundary predictions, ensuring the overall continuity of buildings and the precision of their boundaries. Experiments conducted on the three widely -used building datasets, i.e., Massachusetts, WHU, and Inria, demonstrate that HD -Net achieves the most competitive results with minimal parameter count. In detail, HD -Net outperforms contour -guided and local structure -aware network (CGSANet) with intersection over union (IoU) improvements of 0.40%, 0.95%, and 0.73% on the mentioned datasets, while CGSANet is the state-of-the-art algorithm using a hybrid loss function and deep supervision strategy. Furthermore, the code of the HD -Net will be made available freely at https://github.com/danfenghong/ISPRS\_HD-Net.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
28  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
28  
ISSN:  
0924-2716  
eISSN:  
1872-8235  
入藏号:  
WOS:001178327500001  
Indexed Date:  
2024-03-28  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 100049, Peoples R China  
Institution Address:  
Univ Grenoble Alpes, Inria, CNRS, Grenoble INP,LJK, F-38000 Grenoble, France  
Institution:  
Institution:  
Institution:

### 349. NJUST-CCTD: An Image Database for Milling Tool Wear Classification with

摘要: Deep learning has gained popularity in the task of tool wear identification recently. As an important application of deep learning, however, there exists few public datasets and benchmarks for the research of visual identification of tool wear. To address this issue, we present a classification-based image dataset for carbide milling tool wear (NJUST-CCTD) and make it publicly available on the Github website. This dataset includes two categories: wear tools and no-wear tools. The two categories contain 5000 and 3000 photos, respectively. Based on this dataset, eight baselines are evaluated as references against this benchmark. To further improve the classification performance, we propose a novel cemented carbide milling tool wear intelligent classification framework (CMCNet). The framework consists of two modules: a deep learning based classification network and a multi-scale feature fusion based denoising network called DSSNet. DSSNet is constructed with deeper network structure, connections across layers, and multi-scale sequence fusion module. It is capable of explicitly modeling the semantic and spatial correlation. Apart from DSSNet, the denoising module further improves the performance by adaptively altering the level of denoising based on the performance of the network. The two modules could be optimized with the backward gradient, yielding an end-to-end learning framework. On the basis of the dataset, CMCNet performed exceptionally well when categorizing photos intelligently. After 50 training epochs, the model outperformed the original classification network by 3.250%, achieving the top-1 accuracy of 95.375% on the test set. The NJUST-CCTD can be downloaded at .  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
0268-3768  
eISSN:  
1433-3015  
入藏号:  
WOS:001014988600002  
Indexed Date:  
2023-07-07  
Institution Address:  
Nanjing Univ Sci & Technol, Sch Mech Engn, Nanjing 210094, Jiangsu, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automation, Dept Natl Lab Pattern Recognit, Beijing 100190, Peoples R China  
Institution Address:  
Beijing Res Inst Mech & Elect Technol Ltd, Beijing 100083, Peoples R China  
Institution:  
Institution:  
Institution:  
Beijing Res Inst Mech & Elect Technol Ltd

### 350. Super-resolution Method for Rendered Contents by Multi-scale Feature

摘要: With the development of modern information technology, peoples demand for high resolution and realistic visual perception of image display devices has increased, which has put forward higher requirements for computer software and hardware and brought many challenges to rendering technology in terms of performance and workload. Using machine learning technologies such as deep neural networks to improve the quality and performance of rendered images has become a popular research method in computer graphics, while upsampling low-resolution images through network inference to obtain clearer high-resolution images is an important way to improve image generation performance and ensure high-resolution details. The geometry buffers (G-buffers) generated by the rendering engine in the rendering process contain much semantic information, which help the network learn scene information and features effectively and then improve the quality of upsampling results. In this study, a super-resolution method for rendered contents in low resolution based on deep neural networks is designed. In addition to the color image of the current frame, the method uses high-resolution G-buffers to assist in the calculation and reconstruct the high-resolution content details. The method also leverages a new strategy to fuse the features of highresolution buffers and low-resolution images, which implements a multi-scale fusion of different feature information in a specific fusion module. Experiments demonstrate the effectiveness of the proposed fusion strategy and module, and the proposed method shows obvious advantages, especially in maintaining high-resolution details, when compared with other image super-resolution methods.  
摘要:  
人们对图像显示设备高分辨率和逼真视觉感知的需求随着现代信息技术的发展日益增长,这对计算机软硬件提出了更高要求,也为渲染技术在性能与工作负载上带来更多挑战.利用深度神经网络等机器学习技术对渲染图像进行质量改进和性能提升成为了计算机图形学热门的研究方向,其中通过网络推理将低分辨率图像进行上采样获得更加清晰的高分辨率图像是提升图像生成性能并保证高清细节的一个重要途径.而渲染引擎在渲染流程中产生的几何缓存(geometry buffer, G-buffer)包含较多的语义信息,能够帮助网络有效地学习场景信息与特征,从而提升上采样结果的质量.设计一个基于深度神经网络的低分辨率渲染内容的超分方法.除了当前帧的颜色图像,其使用高分辨率的几何缓存来辅助计算并重建超分后的内容细节.所提方法引入一种新的策略来融合高清缓存与低清图像的特征信息,在特定的融合模块中对不同种特征信息进行多尺度融合.实验验证所提出的融合策略和模块的有效性,并且,在和其他图像超分辨率方法的对比中,所提方法体现出明显的优势,尤其是在高清细节保持方面.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1000-9825  
入藏号:  
CSCD:7736421  
Indexed Date:  
2024-08-15  
Institution Address:  
南京大学, 计算机软件新技术国家重点实验室, 南京, 江苏 210023, 中国  
Institution Address:  
Nanjing University, State Key Laboratory for Novel Software Technology, Nanjing, Jiangsu 210023, China  
Institution:  
南京大学  
Institution:  
Nanjing University

### 351. MShNet: Multi-scale feature combined with h-network for medical image

摘要: Objective: Medical image segmentation is the key foundation of medical image analysis. However, the uncertainty of the size, shape and location of the lesion greatly affect the segmentation accuracy. To solve the above problems, the multi-scale feature combined with the h-network (MShNet) is proposed in this paper. Methods: Firstly, a network framework, which is similar in shape to the letter "h"and consists of an encoder and two decoders is built to obtain stronger feature expression ability. The first decoder is responsible for obtaining the preliminary segmentation information of the image, and the second decoder enhances the feature expression of the nodule by fusing the information learned by the first decoder. Secondly, an enhanced down-sampling module is constructed in the encoder to reduce the information loss caused by down-sampling. In addition, to further reinforce the generalization ability of the model, the fusion convolutional pyramid pooling is designed to realize multi-scale feature fusion. Results: In the internal dataset of thyroid nodules, the DSC is 0.8721 and the HD is only 0.9356; DSC in the public dataset (DDTI,TN3K,ISIC and BUSI) also reached the optimal levels of 0.7580, 0.7815, 0.8853 and 0.7501 respectively and the HD for the last segmentation (Kvasir-SEG) is 16.5197.Conclusion: A large number of experimental results show that MShNet effectively improves the segmen-tation performance with less parameters, and achieves the most advanced performance in robustness and efficiency.Significance: The proposed algorithm provides a deep learning segmentation procedure that can segment thyroid nodule in ultrasound images effectively and efficiently.  
在 BIOSIS Citation Index 中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
被引频次合计:  
13  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:000862722500001  
Indexed Date:  
2022-10-12  
Institution Address:  
Shandong Univ Sci & Technol, Qingdao 266590, Peoples R China  
Institution:

### 352. MSFF-Net: Multi-Scale Feature Fusion Network for Gastrointestinal Vessel

摘要: Purpose The accurate and automatic segmentation of gastrointestinal wall vessels can help to prevent endoscope tip related perforation. Methods based on deep learning and convolution neural network in many different kinds of medical image segmentation tasks have achieved remarkable performance, but in the gastrointestinal vessels segmentation task, because fold and vascular structure characteristics are very similar, blood vessels, fuzzy boundaries, flare and other interference factors, it is very easy to produce false segmentation and rupture of blood vessels. We therefore propose a new multi-scale future fusion network to tackle the aforementioned issues. Methods Our proposed segmentation network consists of encoding, decoding modules, attention module and future fusion module. Through the convolution operation of future fusion module, the output features of each encoder are effectively fused, and the multi-scale information is fully utilized. In addition, we further improve the loss function and enhance the ability of the network to distinguish folds and vessels and predict vascular connectivity by giving different weights to the front background. Results The proposed network is evaluated on our own gastrointestinal wall vessel data set. Experimental results show that compared with the existing advanced vascular segmentation networks, the proposed network has better segmentation performance in the gastrointestinal wall vascular dataset. Conclusion The proposed future fusion method and attention structure loss can better perform feature extraction and fusion according to the characteristics of the gastrointestinal wall vessels to achieve better results.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1609-0985  
eISSN:  
2199-4757  
入藏号:  
WOS:000784941500001  
Indexed Date:  
2022-04-29  
Institution Address:  
Zhejiang Univ Technol, Coll Informat Engn, Hangzhou 310023, Zhejiang, Peoples R China  
Institution Address:  
Zhejiang Univ, Affiliated Hosp 2, Sch Med, Hangzhou 310014, Zhejiang, Peoples R China  
Institution:  
Institution:

### 353. Change Detection for High-resolution Remote Sensing Images Based on a

摘要: The use of bi-temporal remote sensing images for detecting changes in land cover is an important means of obtaining surface change information, thus contributing to urban governance and ecological environment monitoring. In this article, we propose a deep learning model named Siam-UNet3+ for high-resolution remote sensing image change detection. This model integrates the full-scale skip connections and full-scale deep supervision of the network UNet3+, which can achieve the multi-scale feature fusion of remote sensing images, effectively avoiding the locality disadvantage of convolution operations. Different from UNet3+, Siam-UNet3+ has made major improvements, including the following: (1) incorporating a Siamese network in the encoder, which can process bi-temporal remote sensing images in parallel; (2) leveraging the residual module as the backbone, which can avoid gradient vanishing (or exploding) and model degradation problems; (3) adding a Triplet Attention module to the decoder, which can avoid information redundancy that may occur in full-scale skip connections and increase the ability to focus on changing patterns; and (4) designing a hybrid loss function consisting of focal loss and dice loss, which is more suitable for remote sensing image change detection tasks. In this study, we conducted change detection experiments using the publicly available LEVIR-CD dataset, as well as two local datasets in Beijing. Through comparative experiments with five other models and ablation experiments, the proposed model Siam-UNet3+ in this article demonstrated significant advantages and improvements in four evaluation metrics, namely, precision, recall, F1-score, and overall accuracy (OA), proving to have great potential in the application to high-resolution remote sensing image change detection tasks.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0914-4935  
入藏号:  
WOS:001344580300001  
Indexed Date:  
2024-11-04  
Institution Address:  
Beijing Inst Surveying & Mapping, 60 Nanlishi Rd, Beijing 100045, Peoples R China  
Institution Address:  
Beijing Key Lab Urban Spatial Informat Engn, 60 Nanlishi Rd, Beijing 100045, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Remote Sensing & Geomat Engn, 219 Ningliu Rd, Nanjing, Peoples R China  
Institution:  
Beijing Inst Surveying & Mapping  
Institution:  
Beijing Key Lab Urban Spatial Informat Engn  
Institution:

### 354. Object Detector with Residual Learning and Multi-scale Feature

摘要: At present,deep learning has achieved great success in the field of computer vision,but small object detection is still a challenging problem in the field of object detection.Aiming at the problems of low resolution of small objects,blurred images,and less information carried,one object detector that introduces residual learning and multi-scale feature enhancement is proposed.Firstly,an enhanced feature mapping block based on residual learning is introduced into the backbone network.Through channel averaging and normalization,the model more focuses on the object area instead of the background,and it provides additional semantics information for the effective feature layer while taking into account the detection speed.Then the feature map increases the receptive field of the effective feature map through feature fusion block sensitive to context information,and fuses the shallow feature layer and the deep feature layer used for prediction to improve the detection performance at low resolution.Finally,a dual attention block is used to suppress background noise,and key features are embedded in attention.While preserving spatial information,it strengthens the information association between channels,thereby enhancing the expressive ability of features.In order to better detect small objects,the number of a priori boxes for shallow feature mapping is also adjusted.Experimental results show that on the dataset of PASCAL VOC2007,the detection accuracy (mAP) of the algorithm for 300\*300 input scale is 79.9%,which is 2.7 percentage points higher than that of SSD,and the detection accuracy of small objects bird,bottle,chair,and plant is improved 5.1 percentage points,7.5 percentage points,3.9 percentage points,7.2 percentage points,respectively.The detection accuracy (mAP) on the OAP self-made aerial dataset is 82.7%.  
摘要:  
目前深度学习在计算机视觉领域中取得了巨大成功,但是小目标检测仍是目标检测领域中具有挑战性的难题。针对小物体分辨率低、图像模糊、携带信息少等问题,提出了引入残差学习与多尺度特征增强的目标检测器。首先在主干网络中引入基于残差学习的增强特征映射块,通过通道平均和归一化处理使得模型更加专注于对象区域而不是背景,并在兼顾检测速度的同时为有效特征层提供额外的语义信息;然后特征映射对上下文信息敏感的特征融合块进一步增大有效特征图的感受野,并将用于预测的浅特征层与深特征层进行融合,提高低分辨率下的检测性能;最后通过双重注意力块抑制背景噪音,将关键特征嵌入到注意力中,在保留空间信息的同时加强通道间的信息关联,进而增强特征的表达能力。为了更好地检测小目标,还对浅层特征映射先验框数量进行了调整。实验结果表明,在PASCAL VOC2007的数据集上,该算法对于300\*300输入尺度的检测精度(mAP)为79.9%,较SSD提高了2.7个百分点,对小目标bird、bottle、chair、plant检测精度分别提升了5.1个百分点、7.5个百分点、3.9个百分点、7.2个百分点。在OAP自制航拍数据集上的检测精度(mAP)为82.7%。  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
12  
被引频次合计:  
13  
ISSN:  
1673-9418  
入藏号:  
CSCD:7460288  
Indexed Date:  
2023-08-12  
Institution Address:  
江南大学物联网工程学院, 物联网技术应用教育部工程研究中心, 无锡, 江苏 214122, 中国  
Institution Address:  
台州市质量安全检测研究院, 台州, 浙江 318020, 中国  
Institution Address:  
School of Internet of Things Engineering,Jiangnan University, Engineering Research Center of Internet of Things Technology Applications,Ministry of Education, Wuxi, Jiangsu 214122, China  
Institution Address:  
Taizhou Institute of Quality and Safety Testing, Taizhou, Zhejiang 318020, China  
Institution:  
江南大学物联网工程学院  
Institution:  
台州市质量安全检测研究院  
Institution:  
School of Internet of Things Engineering,Jiangnan University  
Institution:  
Taizhou Institute of Quality and Safety Testing

### 355. YOLO-RS: A More Accurate and Faster Object Detection Method for Remote

摘要: In recent years, object detection based on deep learning has been widely applied and developed. When using object detection methods to process remote sensing images, the trade-off between the speed and accuracy of models is necessary, because remote sensing images pose additional difficulties such as complex backgrounds, small objects, and dense distribution to the detection task. This paper proposes YOLO-RS, an optimized object detection algorithm based on YOLOv4 to address the challenges. The Adaptively Spatial Feature Fusion (ASFF) structure is introduced after the feature enhancement network of YOLOv4. It assigns adaptive weight parameters to fuse multi-scale feature information, improving detection accuracy. Furthermore, optimizations are applied to the Spatial Pyramid Pooling (SPP) structure in YOLOv4. By incorporating residual connections and employing 1 x 1 convolutions after maximum pooling, both computation complexity and detection accuracy are improved. To enhance detection speed, Lightnet is introduced, inspired by Depthwise Separable Convolution for reducing model complexity. Additionally, the loss function in YOLOv4 is optimized by introducing the Intersection over Union loss function. This change replaces the aspect ratio loss term with the edge length loss, enhancing sensitivity to width and height, accelerating model convergence, and improving regression accuracy for detected frames. The mean Average Precision (mAP) values of the YOLO-RS model are 87.73% and 92.81% under the TGRS-HRRSD dataset and RSOD dataset, respectively, which are experimentally verified to be 2.15% and 1.66% higher compared to the original YOLOv4 algorithm. The detection speed reached 43.45 FPS and 43.68 FPS, respectively, with 5.29 Frames Per Second (FPS) and 5.30 FPS improvement.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
eISSN:  
2072-4292  
入藏号:  
WOS:001045482300001  
Indexed Date:  
2023-08-20  
Institution Address:  
Nanjing Forestry Univ, Coll Informat Sci & Technol, Nanjing 210037, Peoples R China  
Institution:

### 356. MUFFIN: multi-scale feature fusion for drug-drug interaction prediction

摘要: Motivation: Adverse drug-drug interactions (DDIs) are crucial for drug research and mainly cause morbidity and mortality. Thus, the identification of potential DDIs is essential for doctors, patients and the society. Existing traditional machine learning models rely heavily on handcraft features and lack generalization. Recently, the deep learning approaches that can automatically learn drug features from the molecular graph or drug-related network have improved the ability of computational models to predict unknown DDIs. However, previous works utilized large labeled data and merely considered the structure or sequence information of drugs without considering the relations or topological information between drug and other biomedical objects (e.g. gene, disease and pathway), or considered knowledge graph (KG) without considering the information from the drug molecular structure.  
Results: Accordingly, to effectively explore the joint effect of drug molecular structure and semantic information of drugs in knowledge graph for DDI prediction, we propose a multi-scale feature fusion deep learning model named MUFFIN. MUFFIN can jointly learn the drug representation based on both the drug-self structure information and the KG with rich bio-medical information. In MUFFIN, we designed a bi-level cross strategy that includes cross- and scalar-level components to fuse multi-modal features well. MUFFIN can alleviate the restriction of limited labeled data on deep learning models by crossing the features learned from large-scale KG and drug molecular graph. We evaluated our approach on three datasets and three different tasks including binary-class, multi-class and multi-label DDI prediction tasks. The results showed that MUFFIN outperformed other state-of-the-art baselines.  
Web of Science 核心合集中的 "被引频次":  
125  
在中国科学引文数据库中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
68  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
125  
ISSN:  
1367-4803  
eISSN:  
1367-4811  
入藏号:  
WOS:000697377500017  
PubMed ID:  
33720331  
Indexed Date:  
2021-09-29  
Institution Address:  
Hunan Univ, Sch Comp Sci & Engn, Changsha 410012, Peoples R China  
Institution:

### 357. Region-to-boundary deep learning model with multi-scale feature fusion

摘要: Accurately locating and segmenting lesions, organs, and tissues from medical images are necessary prerequisites for disease diagnosis, monitoring, and treatment planning. Semantic segmentation refers to the classification of each pixel/voxel in two-dimensional or three-dimensional space, which is beneficial to clinical parameter measurement and disease diagnosis. Due to the diversity of features such as size, shape, location, and intensity, segmenting lesions or organs from medical images has always been a challenging worldwide topic. Especially for low-contrast medical images, boundary recognition is particularly difficult. In this paper, we propose a novel region-to-boundary deep learning model to provide a feasible solution to alleviate this problem. First, we use a Ushaped network with two branches behind the last layer, one of which generates the target probability map, and the other obtains the corresponding signed distance map. Secondly, with the help of the signed distance map and obtained multi-scale features, we focus on the boundary of the target lesions or organs to be segmented. Finally, we fuse the region and boundary features and acquire the final results. We conduct extensive experiments on two public data sets and compare with seven the representative methods. The results show that the proposed model is superior to the comparative methods in most evaluation metrics, especially boundary tracking.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
58  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
12  
被引频次合计:  
60  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:000713675900014  
Indexed Date:  
2021-09-17  
Institution Address:  
Hunan Univ, Coll Comp Sci & Elect Engn, Changsha, Peoples R China  
Institution Address:  
Hunan Univ Finance & Econ, Dept Informat Management, Changsha, Peoples R China  
Institution Address:  
SUNY Coll New Paltz, Dept Comp Sci, New Paltz, NY 12561 USA  
Institution:  
Institution:  
Institution:

### 358. Fault-Seg-Net: A method for seismic fault segmentation based on

摘要: Fault identification has important geological significance and practical production value. Due to the effects of earth filtering and environmental noise, it is difficult to identify minor faults, and manual fault identification is inefficient. In this study, an end-to-end deep learning semantic segmentation network Fault-Seg-Net is proposed to identify fault on seismic images, which simultaneously learns global semantic features and local detailed features. In Fault-Seg-Net, a multi-scale residual module is designed to expand the receptive field to mine fine-grained fault features from the low-dimensional feature space. Fault-Seg-Attention module is designed to model long-distance dependencies of pixel spatial location to compensate for the spatial continuity loss. In addition, a compound loss is used to guide the model training to handle imbalanced seismic image segmentation tasks. Experimental results on synthetic datasets have verified that Fault-Seg-Net can achieve high Precision (88.6%), Recall (89.2%), Dice (88.8%) and mIoU (81.5%) simultaneously, which is significantly better than traditional image processing methods and deep learning semantic segmentation networks. Experimental results on real large-scale field datasets have verified that Fault-Seg-Net has important practical value and strong robustness. This study provides an effective solution for intelligent seismic fault identification under complex geological environment.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
21  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
21  
ISSN:  
0266-352X  
eISSN:  
1873-7633  
入藏号:  
WOS:000959588500001  
Indexed Date:  
2023-04-22  
Institution Address:  
China Univ Petr Huadong, Qingdao Inst Software, Coll Comp Sci & Technol, Qingdao, Peoples R China  
Institution:

### 359. A Scalable Target Orientation Detection Method for Remote Sensing Images

摘要: Significant progress has been achieved in the development of oriented target detection algorithms based on deep learning, which have found widespread application in various fields, including remote sensing. However, existing methods struggle with adjusting model size and often exhibit unsatisfactory detection performance for targets that overlap, are large, or have similar backgrounds. To address these challenges, this letter proposes an oriented target detection algorithm called Oriented you only look once X (YOLOX), which integrates several optimization techniques. Specifically, to meet the requirements of oriented detection while enhancing feature extraction, we introduce a new network architecture that includes an orientation detection branch and a multiscale feature fusion module (MSFFM). An MSFFM based on attention weights is proposed to integrate features across scales while minimizing noise. In addition, to mitigate the impact of the number of positive samples on the original loss function and focus the network's attention on learning challenging targets, an object-aware reweighted loss function is introduced in this study. This approach dynamically adjusts the loss contribution for each target. Two models of different sizes are developed using the Oriented YOLOX scaling strategy to cater to scenarios prioritizing either accuracy or speed. Extensive experiments on the dataset for object detection in aerial images (DOTA) and object detection in optical remote sensing images (DIOR-R) datasets demonstrate that Oriented YOLOX performs better in detecting challenging targets. Compared with other oriented target detection methods, this approach not only achieves higher detection accuracy but also reduces parameter counts, improving inference speed.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001313341800007  
Indexed Date:  
2024-09-21  
Institution Address:  
Xidian Univ, Sch Artificial Intelligence, Xian 710071, Peoples R China  
Institution Address:  
China Elect Technol Grp Corp, Res Inst 54, Shijiazhuang 050081, Peoples R China  
Institution:  
Institution:

### 360. Rainy day image semantic segmentation based on two-stage progressive

摘要: Semantic segmentation plays a crucial role in the fields of computer vision and computer graphics, with extensive applications in various practical scenarios. Significant progress has been made in semantic segmentation tasks using deep learning-based methods. However, most existing semantic segmentation algorithms focus on good weather conditions, and they face challenges in terms of accuracy and robustness when applied to rainy scenes due to factors such as raindrops, haze, and lighting variations. To address this issue, this paper proposes a rainy-day semantic segmentation method based on a two-stage progressive network. The proposed method consists of two modules: a rain removal module responsible for eliminating raindrops and haze from the input rainy images and restoring the basic structural information of the images, and a segmentation module that performs pixel-level semantic prediction on the rain-removed images. Specifically, the rain removal module introduces two progressive units with shared weights to gradually achieve rain removal. The segmentation module adopts an encoder-decoder architecture, utilizing down-sampling and deep asynchronous bottleneck units for encoding. It also introduces a dual attention-guided fusion module to aggregate channel attention information and spatial attention information, guiding the multiscale feature fusion process in the decoder. Experimental results demonstrate that this method effectively mitigates the influence of rain streaks on semantic segmentation, thereby improving segmentation performance and achieving more accurate and robust semantic segmentation results in rainy conditions. We will provide the code and datasets on https://github.com/zhang152267/TSPN.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001170555600001  
Indexed Date:  
2024-03-09  
Institution Address:  
Hebei Univ Engn, Sch Informat & Elect Engn, Handan, Peoples R China  
Institution:

### 361. Self-Paced Feature Attention Fusion Network for Concealed Object

摘要: The active millimeter-wave (AMMW) scanner has been widely used for inspecting human security in public places in recent years owing to its ability to detect all kinds of objects under the clothes and be harmless to the body. However, it is really challenging to detect all concealed objects automatically and accurately due to inherent imaging noise, unknown object kind, and uncertain position. Recently, many existing methods, especially deep learning-based, have achieved good performances on concealed object detection. These methods work well for detecting a few kinds of large objects, but fail to perform on dim and incomplete hard objects. To address this task, a concealed object detection model with self-paced feature attention fusion network (SPFAFN) is proposed in this article. To be specific, the features with different scales are fused in a top-down manner to integrate details and global semantics to better detect small objects. During fusing multi-scale features, a hierarchical pyramid attention mechanism composed of channel and spatial attention is developed to perceive the object. Moreover, boosting self-paced learning is exploited to guide the model to learn hard samples that are difficultly detected. The proposed method is validated on two real-world datasets: an AMMW dataset and a publicly available passive millimeter-wave (PMMW) dataset. Experimental results demonstrate that the proposed approach is superior to the state-of-the-art methods, and achieves better performances on the two datasets with Average Precision (AP).  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
28  
被引频次合计:  
29  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:000742183600022  
Indexed Date:  
2022-01-21  
Institution Address:  
Xidian Univ, Sch Artificial Intelligence, Xian 710071, Peoples R China  
Institution Address:  
Beijing Huahang Radio Measurement Res Inst, Beijing 100013, Peoples R China  
Institution:  
Institution:  
Beijing Huahang Radio Measurement Res Inst

### 362. New Fusion Network with Dual-Branch Encoder and Triple-Branch Decoder

摘要: Deep learning plays a highly essential role in the domain of remote sensing change detection (CD) due to its high efficiency. From some existing methods, we can observe that the fusion of information at each scale is quite vital for the accuracy of the CD results, especially for the common problems of pseudo-change and the difficult detection of change edges in the CD task. With this in mind, we propose a New Fusion network with Dual-branch Encoder and Triple-branch Decoder (DETDNet) that follows a codec structure as a whole, where the encoder adopts a siamese Res2Net-50 structure to extract the local features of the bitemporal images. As for the decoder in previous works, they usually employed a single branch, and this approach only preserved the fusion features of the encoder's bitemporal images. Distinguished from these approaches, we adopt the triple-branch architecture in the decoder for the first time. The triple-branch structure preserves not only the dual-branch features from the encoder in the left and right branches, respectively, to learn the effective and powerful individual features of each temporal image but also the fusion features from the encoder in the middle branch. The middle branch utilizes triple-branch aggregation (TA) to realize the feature interaction of the three branches in the decoder, which enhances the integrated features and provides abundant and supplementary bitemporal feature information to improve the CD performance. The triple-branch architecture of the decoder ensures that the respective features of the bitemporal images as well as their fused features are preserved, making the feature extraction more integrated. In addition, the three branches employ a multiscale feature extraction module (MFE) per layer to extract multiscale contextual information and enhance the feature representation capability of the CD. We conducted comparison experiments on the BCDD, LEVIR-CD, and SYSU-CD datasets, which were created in New Zealand, the USA, and Hong Kong, respectively. The data were preprocessed to contain 7434, 10,192, and 20,000 image pairs, respectively. The experimental results show that DETDNet achieves F1 scores of 92.7%, 90.99%, and 81.13%, respectively, which shows better results compared to some recent works, which means that the model is more robust. In addition, the lower FP and FN indicate lower error and misdetection rates. Moreover, from the analysis of the experimental results, compared with some existing methods, the problem of pseudo-changes and the difficulty of detecting small change areas is better solved.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:000994366700001  
Indexed Date:  
2023-06-10  
Institution Address:  
Xinjiang Univ, Coll Informat Sci & Engn, Urumqi 830046, Peoples R China  
Institution:

### 363. 3D Attention M-net for Short-axis Left Ventricular Myocardium

摘要: Small rodent cardiac magnetic resonance imaging (MRI) plays an important role in preclinical models of cardiac disease. Accurate myocardial boundaries delineation is crucial to most morphological and functional analysis in rodent cardiac MRIs. However, rodent cardiac MRIs, due to animal's small cardiac volume and high heart rate, are usually acquired with sub-optimal resolution and low signal-to-noise ratio (SNR). These rodent cardiac MRIs can also suffer from signal loss due to the intra-voxel dephasing. These factors make automatic myocardial segmentation challenging. Manual contouring could be applied to label myocardial boundaries but it is usually laborious, time consuming, and not systematically objective. In this study, we present a deep learning approach based on 3D attention M-net to perform automatic segmentation of left ventricular myocardium. In the deep learning architecture, we use dual spatial-channel attention gates between encoder and decoder along with multi-scale feature fusion path after decoder. Attention gates enable networks to focus on relevant spatial information and channel features to improve segmentation performance. A distance derived loss term, besides general dice loss and binary cross entropy loss, was also introduced to our hybrid loss functions to refine segmentation contours. The proposed model outperforms other generic models, like U-Net and FCN, in major segmentation metrics including the dice score (0.9072), Jaccard index (0.8307) and Hausdorff distance (3.1754 pixels), which are comparable to the results achieved by state-of-the-art models on human cardiac ACDC17 datasets.  
会议名称:  
43rd Annual International Conference of the IEEE-Engineering-in-Medicine-and-Biology-Society (IEEE EMBC)  
会议日期:  
NOV 01-05, 2021  
会议地点:  
ELECTR NETWORK  
会议赞助商:  
IEEE Engn Med & Biol Soc; IEEE; Elsevier; Inst Engn & Technol  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1557-170X  
eISSN:  
1558-4615  
ISBN:  
978-1-7281-1179-7  
入藏号:  
WOS:000760910503068  
PubMed ID:  
34891958  
Indexed Date:  
2022-04-21  
Institution Address:  
Johns Hopkins Univ, Ctr Imaging Sci, Baltimore, MD 21218 USA  
Institution Address:  
Johns Hopkins Med Inst, Sect Cardiol, Dept Med, Baltimore, MD 21218 USA  
Institution:  
Institution:

### 364. Detecting XSS with Random Forest and Multi-Channel Feature Extraction

摘要: In the era of the Internet, widely used web applications have become the target of hacker attacks because they contain a large amount of personal information. Among these vulnerabilities, stealing private data through cross- site scripting (XSS) attacks is one of the most commonly used attacks by hackers. Currently, deep learning-based XSS attack detection methods have good application prospects; however, they suffer from problems such as being prone to overfitting, a high false alarm rate, and low accuracy. To address these issues, we propose a multi-stage feature extraction and fusion model for XSS detection based on Random Forest feature enhancement. The model utilizes Random Forests to capture the intrinsic structure and patterns of the data by extracting leaf node indices as features, which are subsequently merged with the original data features to form a feature set with richer information content. Further feature extraction is conducted through three parallel channels. Channel I utilizes parallel onedimensional convolutional layers (1D convolutional layers) with different convolutional kernel sizes to extract local features at different scales and perform multi-scale feature fusion; Channel II employs maximum one-dimensional pooling layers (max 1D pooling layers) of various sizes to extract key features from the data; and Channel III extracts global information bi-directionally using a Bi-Directional Long-Short Term Memory Network (Bi-LSTM) and incorporates a multi-head attention mechanism to enhance global features. Finally, effective classification and prediction of XSS are performed by fusing the features of the three channels. To test the effectiveness of the model, we conduct experiments on six datasets. We achieve an accuracy of 100% on the UNSW-NB15 dataset and 99.99% on the CICIDS2017 dataset, which is higher than that of the existing models.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001290839400002  
Indexed Date:  
2024-08-21  
Institution Address:  
Beijing Union Univ, Smart City Coll, Beijing 100101, Peoples R China  
Institution:

### 365. Multi-Resolution Supervision Network with an Adaptive Weighted Loss for

摘要: Desert segmentation of remote sensing images is the basis of analysis of desert area. Desert images are usually characterized by large image size, large-scale change, and irregular location distribution of surface objects. The multi-scale fusion method is widely used in the existing deep learning segmentation models to solve the above problems. Based on the idea of multi-scale feature extraction, this paper took the segmentation results of each scale as an independent optimization task and proposed a multi-resolution supervision network (MrsSeg) to further improve the desert segmentation result. Due to the different optimization difficulty of each branch task, we also proposed an auxiliary adaptive weighted loss function (AWL) to automatically optimize the training process. MrsSeg first used a lightweight backbone to extract different-resolution features, then adopted a multi-resolution fusion module to fuse the local information and global information, and finally, a multi-level fusion decoder was used to aggregate and merge the features at different levels to get the desert segmentation result. In this method, each branch loss was treated as an independent task, AWL was proposed to calculate and adjust the weight of each branch. By giving priority to the easy tasks, the improved loss function could effectively improve the convergence speed of the model and the desert segmentation result. The experimental results showed that MrsSeg-AWL effectively improved the learning ability of the model and has faster convergence speed, lower parameter complexity, and more accurate segmentation results.  
Web of Science 核心合集中的 "被引频次":  
7  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
eISSN:  
2072-4292  
入藏号:  
WOS:000660615900001  
Indexed Date:  
2021-06-21  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Jiangsu Collaborat Innovat Ctr Atmospher Environm, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Forestry Univ, Coll Informat Sci & Technol, Nanjing 210037, Peoples R China  
Institution:  
Institution:

### 366. A Hybrid Efficient U-Net Framework for Detection of Anterior Belly of

摘要: The digastric muscle is an important muscle involved in functions such as chewing and swallowing. Ultrasonography is the preferred method for imaging the soft tissues of the head and neck but is highly operator-dependent. Artificial intelligence, particularly deep learning-based segmentation models, has the potential to improve the accuracy and precision of ultrasound images. In this study, a MultiResUNet-Fusion model including residual blocks, multiscale feature fusion, and SE blocks was developed for segmentation of the anterior belly of the digastric muscle. The model was trained on 198 ultrasound images from 99 participants. Combo Loss (a combination of Binary Cross-Entropy and Dice Loss) was used to train the model and segmentation metrics such as F1-score, Intersection over Union (IoU) and Dice Co-efficient were used to evaluate performance. The proposed MultiResUNet-Fusion model provided high accuracy and reliability for the segmentation of the anterior belly of the digastric muscle. The proposed MultiResUNet-Fusion model demonstrated high performance by achieving F1 score (95.38%) and IoU (91.17%). The visual results showed that the segmentation masks of the MultiResUNet-Fusion models provided predictions close to the real labels, and all models generally localized the region of interest accurately. The MultiResUNet-Fusion model provides high accuracy in low-contrast ultrasound images, making it suitable for clinical applications. The model can contribute to clinical diagnostic processes with its ability to accurately detect small and large structures. Future studies can increase the generalization capacity of the model by testing it in different modalities.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001410255600023  
Indexed Date:  
2025-02-09  
Institution Address:  
Giresun Univ, Fac Dent, Dept Dentomaxillofacial Radiol, TR-28200 Giresun, Turkiye  
Institution Address:  
Samsun Univ, Fac Engn, Dept Software Engn, TR-55420 Samsun, Turkiye  
Institution Address:  
Firat Univ, Fac Technol, Dept Elect & Elect Engn, TR-23119 Elazig, Turkiye  
Institution Address:  
King Abdulaziz Univ, Fac Engn, Dept Elect & Elect Engn, Jeddah 21589, Saudi Arabia  
Institution:  
Institution:  
Institution:  
Institution:

### 367. Real-time instance segmentation of surgical instruments using attention

摘要: Precise instrument segmentation aids surgeons to navigate the body more easily and increases patient safety. While accurate tracking of surgical instruments in real-time plays a crucial role in minimally invasive computer -assisted surgeries, it is a challenging task to achieve, mainly due to: (1) a complex surgical environment, and (2) model design trade-off in terms of both optimal accuracy and speed. Deep learning gives us the opportunity to learn complex environment from large surgery scene environments and placements of these instruments in real world scenarios. The Robust Medical Instrument Segmentation 2019 challenge (ROBUST-MIS) provides more than 10,000 frames with surgical tools in different clinical settings. In this paper, we propose a light-weight single stage instance segmentation model complemented with a convolutional block attention module for achieving both faster and accurate inference. We further improve accuracy through data augmentation and optimal anchor localization strategies. To our knowledge, this is the first work that explicitly focuses on both real-time performance and improved accuracy. Our approach out-performed top team performances in the most recent edition of ROBUST-MIS challenge with over 44% improvement on area-based multi-instance dice metric MI\_DSC and 39% on distance-based multi-instance normalized surface dice MI\_NSD. We also demonstrate real-time performance (> 60 frames-per-second) with different but competitive variants of our final approach.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
23  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
23  
ISSN:  
1361-8415  
eISSN:  
1361-8423  
入藏号:  
WOS:000847945700004  
PubMed ID:  
35985195  
Indexed Date:  
2022-09-07  
Institution Address:  
Tecnol Monterrey, Escuela Ingn & Ciencias, Monterrey, NL, Mexico  
Institution Address:  
Univ Oxford, Inst Biomed Engn IBME, Dept Engn Sci, Oxford, England  
Institution Address:  
Univ Oxford, Oxford NIHR Biomed Res Ctr, Oxford, England  
Institution Address:  
Univ Leeds, Sch Comp, Leeds, W Yorkshire, England  
Institution:  
Institution:  
Institution:  
Institution:

### 368. The Cost of Urban Renewal: Annual Construction Waste Estimation via

摘要: Construction waste is an inevitable byproduct of urban renewal, causing severe pressure on the environment, health, and ecology. Accurately estimating the production of construction waste is crucial for assessing the consumption of urban renewal. However, traditional manual estimation methods rely heavily on statistical data and historical experience, which lack flexibility in practical applications and are time-consuming and labor-intensive. In addition, their accuracy and timeliness need to be improved urgently. Fortunately, with the advantages of high-resolution remote sensing images (HRSIs) such as strong timeliness, large amounts of information, and macroscopic observations, they are suitable for the large-scale dynamic change detection of construction waste. However, the existing deep learning models have a relatively poor ability to extract and fuse features for small and multi-scale targets, and it is difficult to deal with irregularly shaped and fragmented detection areas. Therefore, this study proposes a Multi-scale Target Attention-Enhanced Network (MT-AENet), which is used to dynamically track and detect changes in buildings and construction waste disposal sites through HRSIs and accurately estimate the annual production of urban construction waste. The MT-AENet introduces a novel encoder-decoder architecture. In the encoder, ResNet-101 is utilized to extract high-level semantic features. A depthwise separable-atrous spatial pyramid pooling (DS-ASPP) module with different dilation rates is constructed to address insufficient receptive fields, resolving the issue of discontinuous holes when extracting large targets. A dual-attention mechanism module (DAMM) is employed to better preserve positional and channel details. In the decoder, multi-scale feature fusion (MS-FF) is utilized to capture contextual information, integrating shallow and intermediate features of the backbone network, thereby enhancing extraction capabilities in complex scenes. The MT-AENet is used to extract buildings and construction waste at different periods in the study area, and the actual production and landfill volume of construction waste are calculated based on area changes, indirectly measuring the rate of urban construction waste resource conversion. The experimental results in Changping District, Beijing demonstrate that the MT-AENet outperforms existing baseline networks in extracting buildings and construction waste. The results of this study are validated according to government statistical standards, providing a promising direction for efficiently analyzing the consumption of urban renewal.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:001245294700001  
Indexed Date:  
2024-06-21  
Institution Address:  
Beijing Univ Technol, Fac Informat Technol, Beijing 100124, Peoples R China  
Institution Address:  
Chinese Acad Sci, State Key Lab Remote Sensing Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Beijing Big Data Ctr, Beijing 100101, Peoples R China  
Institution Address:  
Jimei Univ, Nav Coll, Xiamen 361021, Peoples R China  
Institution Address:  
Beijing Econ Technol Dev Area, Adm Examinat & Approval Bur, Beijing 100176, Peoples R China  
Institution:  
Institution:  
Institution:  
Beijing Big Data Ctr  
Institution:  
Institution:  
Beijing Econ Technol Dev Area

### 369. Attention-Guided Deep Neural Network With Multi-Scale Feature Fusion for

摘要: Liver vessel segmentation is fast becoming a key instrument in the diagnosis and surgical planning of liver diseases. In clinical practice, liver vessels are normally manual annotated by clinicians on each slice of CT images, which is extremely laborious. Several deep learning methods exist for liver vessel segmentation, however, promoting the performance of segmentation remains a major challenge due to the large variations and complex structure of liver vessels. Previous methods mainly using existing UNet architecture, but not all features of the encoder are useful for segmentation and some even cause interferences. To overcome this problem, we propose a novel deep neural network for liver vessel segmentation, called LVSNet, which employs special designs to obtain the accurate structure of the liver vessel. Specifically, we design Attention-Guided Concatenation (AGC) module to adaptively select the useful context features from low-level features guided by high-level features. The proposed AGC module focuses on capturing rich complemented information to obtain more details. In addition, we introduce an innovative multi-scale fusion block by constructing hierarchical residual-like connections within one single residual block, which is of great importance for effectively linking the local blood vessel fragments together. Furthermore, we construct a new dataset containing 40 thin thickness cases (0.625 mm) which consist of CT volumes and annotated vessels. To evaluate the effectiveness of the method with minor vessels, we also propose an automatic stratification method to split major and minor liver vessels. Extensive experimental results demonstrate that the proposed LVSNet outperforms previous methods on liver vessel segmentation datasets. Additionally, we conduct a series of ablation studies that comprehensively support the superiority of the underlying concepts.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
22  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
78  
被引频次合计:  
82  
ISSN:  
2168-2194  
eISSN:  
2168-2208  
入藏号:  
WOS:000678341200027  
PubMed ID:  
33264097  
Indexed Date:  
2021-08-11  
Institution Address:  
Univ Adelaide, Sch Comp Sci, Adelaide, SA 5005, Australia  
Institution Address:  
Tsinghua Univ THU, State Key Lab Precis Measurement Technol & Instru, Dept Precis Instrument, Innovat Ctr Future Chips, Beijing, Peoples R China  
Institution Address:  
Beijing Jingzhen Med Technol Ltd, Beijing, Peoples R China  
Institution Address:  
Tsinghua Univ, State Key Lab Precis Measurement Technol & Instru, Beijing, Peoples R China  
Institution Address:  
Tsinghua Univ, Inst Interdisciplinary Informat Sci, Beijing, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Sch Comp Sci, Xian, Shaanxi, Peoples R China  
Institution Address:  
Xidian Univ, Sch Comp Sci & Technol, Xian, Peoples R China  
Institution:  
Institution:  
Institution:  
Beijing Jingzhen Med Technol Ltd  
Institution:  
Institution:  
Institution:  
Institution:

### 370. TSMAL: Target-Shadow Mask Assistance Learning Network for SAR Target

摘要: Deep learning-based synthetic aperture radar (SAR) target recognition methods mainly emphasize the amplitude characteristics resulting from backscatter at the target's principal scattering points. Shadows, as critical by-products of SAR imaging, encapsulate vital details regarding the target's structural configuration. In order to effectively utilize the target information and shadow information in SAR images, we propose a novel target-shadow mask assistance learning (TSMAL) network for SAR target recognition. It systematically leverages domain knowledge in SAR images through three key points: data preprocessing, network structure, and multitask loss function. Specifically, the data preprocessing, with the help of the segmentation algorithm, extracts the target mask and shadow mask as domain knowledge in the SAR image to be utilized. Then, the target-shadow mask assistance (TSMA) layer is designed to learn complementary representations within each convolutional layer by exploiting target-shadow information. The TSMA enhances the scatter features related to the target regions as well as the shape features contained in the shadow regions, and suppresses backgrounds. Meanwhile, the multilayer coordinate attention (MCA) is used for multiscale feature fusion. Finally, a multitask learning loss is designed depending on the recognition and feature optimization tasks to guide the network learning. By synergistically employing a TSMAL alongside multitask learning strategies, the network proficiently acquires both target and shadow features. The experiments conducted on the moving and stationary target acquisition and recognition (MSTAR) dataset show that the proposed method can effectively enhance the feature extraction ability in the target and shadow regions, and improve the performance of SAR target recognition.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001335560700013  
Indexed Date:  
2024-10-26  
Institution Address:  
Xidian Univ, Natl Key Lab Radar Signal Proc, Xian 710071, Peoples R China  
Institution Address:  
Xidian Univ, Natl Key Lab Radar Signal Proc, Xian 710071, Peoples R China  
Institution Address:  
Xidian Univ, Hangzhou Inst Technol, Hangzhou 311200, Peoples R China  
Institution Address:  
Xidian Univ, Guangzhou Inst Technol, Guangzhou 510555, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 371. MultiR-Net: A Novel Joint Learning Network for COVID-19 segmentation and

摘要: The outbreak of COVID-19 has caused a severe shortage of healthcare resources. Ground Glass Opacity (GGO) and consolidation of chest CT scans have been an essential basis for imaging diagnosis since 2020. The similarity of imaging features between COVID-19 and other pneumonia makes it challenging to distinguish between them and affects radiologists' diagnosis. Recently, deep learning in COVID-19 has been mainly divided into disease classification and lesion segmentation, yet little work has focused on the feature correlation between the two tasks. To address these issues, in this study, we propose MultiR-Net, a 3D deep learning model for combined COVID-19 classification and lesion segmentation, to achieve real-time and interpretable COVID-19 chest CT diagnosis. Precisely, the proposed network consists of two subnets: a multi-scale feature fusion UNet-like subnet for lesion segmentation and a classification subnet for disease diagnosis. The features between the two subnets are fused by the reverse attention mechanism and the iterable training strategy. Meanwhile, we proposed a loss function to enhance the interaction between the two subnets. Individual metrics can not wholly reflect network effectiveness. Thus we quantify the segmentation results with various evaluation metrics such as average surface distance, volume Dice, and test on the dataset. We employ a dataset containing 275 3D CT scans for classifying COVID-19, Community-acquired Pneumonia (CAP), and healthy people and segmented lesions in pneumonia patients. We split the dataset into 70% and 30% for training and testing. Extensive experiments showed that our multi-task model framework obtained an average recall of 93.323%, an average precision of 94.005% on the classification test set, and a 69.95% Volume Dice score on the segmentation test set of our dataset.  
Web of Science 核心合集中的 "被引频次":  
22  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
被引频次合计:  
22  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:000806836400003  
PubMed ID:  
35305504  
Indexed Date:  
2022-06-22  
Institution Address:  
Shanghai Univ, Sch Comp Engn & Sci, Shangda Rd, Shanghai 200444, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Shanghai Peoples Hosp 9, Sch Med, Mohe Rd, Shanghai 200111, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ Affiliated Peoples Hosp 6, Inst Diagnost & Intervent Radiol, Yishan Rd, Shanghai 200233, Peoples R China  
Institution:  
Institution:  
Institution:

### 372. GTCBS-YOLOv5s: A lightweight model for weed species identification in

摘要: Accurate and rapid weed species identification contributes to selective herbicide spraying or robotic weeding. An image-based method for automatic identification of weed species in paddy fields is highly desirable. However, water reflection, soil background, occlusion, and varying growth and illumination make the development of this method challenging. To address this issue, an improved deep learning model GTCBS-YOLOv5s was proposed to identify six weed species in paddy fields, where Ghost, C3Trans and convolutional block attention module (CBAM) were employed to improve weed feature extraction in complex environments. The bidirectional feature pyramid network (BiFPN) coupled with Concat structure was introduced in the Neck network to achieve the multi-scale feature fusion for identifying various weed species. Three different output feature maps in Detect network were utilized to identify weeds of varying sizes. A more comprehensive scale-sensitive intersection over union (SIoU) loss function was adopted to eliminate the redundant generating boxes. The results showed that GTCBS-YOLOv5s model achieved a mean average precision (mAP) of 91.1 % for the test set, and the identification speed reached 85.7 FPS. Robustness tests demonstrated that GTCBS-YOLOv5s obtained satisfactory performance in identifying weeds under various lighting conditions, with precision (P), recall (R) and mAP all greater than 85 %. Occluded weeds were identified with P, R and average precision (AP) greater than 89.8 %, 90.1 % and 90.3 %, respectively. Furthermore, GTCBS-YOLOv5s had good performance in identifying weeds at different growth stages, with P, R and mAP higher than 90.1 %, 89.5 %, and 90.3 % respectively. Compared with the state-of-the-art models, GTCBS-YOLOv5s was highly promising for deployment to the embedded devices for real-time field detection due to its high accuracy, lightweight and robust attributes, as well as fast inference.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:001149442200001  
Indexed Date:  
2024-02-04  
Institution Address:  
Shandong Agr Univ, Coll Mech & Elect Engn, Tai An 271018, Peoples R China  
Institution Address:  
South China Agr Univ, Coll Engn, Guangzhou 510642, Peoples R China  
Institution Address:  
Jining Acad Agr Sci, Jining 272031, Shandong, Peoples R China  
Institution:  
Institution:  
Institution:  
Jining Acad Agr Sci

### 373. OA-MEN: a fusion deep learning approach for enhanced accuracy in knee

摘要: Background Knee osteoarthritis (KOA) constitutes the prevailing manifestation of arthritis. Radiographs function as a common modality for primary screening; however, traditional X-ray evaluation of osteoarthritis confronts challenges such as reduced sensitivity, subjective interpretation, and heightened misdiagnosis rates. The objective of this investigation is to enhance the validation and optimization of accuracy and efficiency in KOA assessment by utilizing fusion deep learning techniques.Methods This study aims to develop a highly accurate and lightweight model for automatically predicting and classifying KOA through knee X-ray imaging. We propose a deep learning model named OA-MEN, which integrates a hybrid model combining ResNet and MobileNet feature extraction with multi-scale feature fusion. This approach ensures enhanced extraction of semantic information without losing the advantages of large feature maps provided by high image resolution in lower layers of the network. This effectively expands the model's receptive field and strengthens its understanding capability. Additionally, we conducted unseen-data tests and compared our model with widely used baseline models to highlight its superiority over conventional approaches.Results The OA-MEN model demonstrated exceptional performance in tests. In the unseen-data test, our model achieved an average accuracy (ACC) of 84.88% and an Area Under the Curve (AUC) of 89.11%, marking improvements over the best-performing baseline models. These results showcase its improved capability in predicting KOA from X-ray images, making it a promising tool for assisting radiologists in diagnosis and treatment selection in clinical settings.Conclusion Leveraging deep learning for osteoarthritis classification guarantees heightened efficiency and accuracy. The future goal is to seamlessly integrate deep learning and advanced computational techniques with the expertise of medical professionals.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2296-4185  
入藏号:  
WOS:001397529200001  
PubMed ID:  
39830688  
Indexed Date:  
2025-01-21  
Institution Address:  
Ningxia Med Univ, Gen Hosp, Dept Radiol, Yinchuan, Peoples R China  
Institution Address:  
Univ Sains Malaysia, Sch Hlth Sci, Kota Baharu, Kelantan, Malaysia  
Institution Address:  
Sichuan Univ, Coll Biomed Engn, Chengdu, Sichuan, Peoples R China  
Institution Address:  
Ningxia Med Univ, Gen Hosp, Dept Orthoped, Yinchuan, Peoples R China  
Institution Address:  
Tsinghua Univ, Sch Clin Med, Beijing, Peoples R China  
Institution Address:  
Beijing Hosp, Natl Ctr Gerontol, Dept Rheumatol & Immunol, Beijing, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 374. A Novel Multi-Scale Feature Map Fusion for Oil Spill Detection of SAR

摘要: The efficient and timely identification of oil spill areas is crucial for ocean environmental protection. Synthetic aperture radar (SAR) is widely used in oil spill detection due to its all-weather monitoring capability. Meanwhile, existing deep learning-based oil spill detection methods mainly rely on the classical U-Net framework and have achieved impressive results. However, SAR images exhibit high noise, blurry boundaries, and irregular shapes of target areas, as well as speckles and shadows, which lead to the loss of performance in existing algorithms. In this paper, we propose a novel network architecture to achieve more precise segmentation of oil spill areas by reintroducing rich semantic contextual information before obtaining the final segmentation mask. Specifically, the proposed architecture can re-fuse feature maps from different levels at the decoder end. We design a multi-convolutional layer (MCL) module to extract basic feature information from SAR images, and a feature extraction module (FEM) module further extracts and fuses feature maps generated by the U-Net decoder at different levels. Through these operations, the network can learn rich global and local contextual information, enable sufficient interaction of feature information at different stages, enhance the model's contextual awareness, and improve its ability to recognize complex textures and blurry boundaries, thereby enhancing the segmentation accuracy of SAR images. Compared to many U-Net based segmentation networks, our method shows promising results and achieves state-of-the-art performance on multiple evaluation metrics.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
7  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
eISSN:  
2072-4292  
入藏号:  
WOS:001231288700001  
Indexed Date:  
2024-05-31  
Institution Address:  
Harbin Inst Technol, Sch Comp Sci & Technol, Weihai 264209, Peoples R China  
Institution Address:  
Guangzhou Univ, Sch Elect & Commun Engn, Guangzhou 510006, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution:  
Institution:  
Institution:

### 375. Modality preserving U-Net for segmentation of multimodal medical images

摘要: Background: Recent advances in artificial intelligence and digital image processing have inspired the use of deep neural networks for segmentation tasks in multimodal medical imaging. Unlike natural images, multimodal medical images contain much richer information regarding different modal properties and therefore present more challenges for semantic segmentation. However, there is no report on systematic research that integrates multi-scaled and structured analysis of single-modal and multimodal medical images.  
Methods: We propose a deep neural network, named as Modality Preserving U-Net (MPU-Net), for modality-preserving analysis and segmentation of medical targets from multimodal medical images. The proposed MPU-Net consists of a modality preservation encoder (MPE) module that preserves the feature independency among the modalities and a modality fusion decoder (MFD) module that performs a multiscale feature fusion analysis for each modality in order to provide a rich feature representation for the final task. The effectiveness of such a single-modal preservation and multimodal fusion feature extraction approach is verified by multimodal segmentation experiments and an ablation study using brain tumor and prostate datasets from Medical Segmentation Decathlon (MSD).  
Results: The segmentation experiments demonstrated the superiority of MPU-Net over other methods in the segmentation tasks for multimodal medical images. In the brain tumor segmentation tasks, the Dice scores (DSCs) for the whole tumor (WT), the tumor core (TC) and the enhancing tumor (ET) regions were 89.42%, 86.92%, and 84.59%, respectively. In the meanwhile, the 95% Hausdorff distance (HD95) results were 3.530, 4.899 and 2.555, respectively. In the prostate segmentation tasks, the DSCs for the peripheral zone (PZ) and the transitional zone ( TZ) of the prostate were 71.20% and 90.38%, respectively. In the meanwhile, the 95% HD95 results were 6.367 and 4.766, respectively. The ablation study showed that the combination of single-modal preservation and multimodal fusion methods improved the performance of multimodal medical image feature analysis.  
Conclusions: In the segmentation tasks using brain tumor and prostate datasets, the MPU-Net method has achieved the improved performance in comparison with the conventional methods, indicating its potential application for other segmentation tasks in multimodal medical images.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
在 BIOSIS Citation Index 中的被引频次:  
2  
被引频次合计:  
9  
ISSN:  
2223-4292  
eISSN:  
2223-4306  
入藏号:  
WOS:001011927300001  
PubMed ID:  
37581055  
Indexed Date:  
2023-07-02  
Institution Address:  
Univ Sci & Technol China, Dept Precis Machinery & Precis Instrumentat, Hefei, Peoples R China  
Institution Address:  
Univ Sci & Technol China, Suzhou Inst Adv Res, 188 Renai Rd, Suzhou 215123, Peoples R China  
Institution Address:  
Univ Sci & Technol China, Sch Microelect, Hefei, Peoples R China  
Institution:  
Institution:  
Institution:

### 376. MSAt-GAN: a generative adversarial network based on multi-scale and deep

摘要: For the past few years, image fusion technology has made great progress, especially in infrared and visible light image infusion. However, the fusion methods, based on traditional or deep learning technology, have some disadvantages such as unobvious structure or texture detail loss. In this regard, a novel generative adversarial network named MSAt-GAN is proposed in this paper. It is based on multi-scale feature transfer and deep attention mechanism feature fusion, and used for infrared and visible image fusion. First, this paper employs three different receptive fields to extract the multi-scale and multi-level deep features of multi-modality images in three channels rather than artificially setting a single receptive field. In this way, the important features of the source image can be better obtained from different receptive fields and angles, and the extracted feature representation is also more flexible and diverse. Second, a multi-scale deep attention fusion mechanism is designed in this essay. It describes the important representation of multi-level receptive field extraction features through both spatial and channel attention and merges them according to the level of attention. Doing so can lay more emphasis on the attention feature map and extract significant features of multi-modality images, which eliminates noise to some extent. Third, the concatenate operation of the multi-level deep features in the encoder and the deep features in the decoder are cascaded to enhance the feature transmission while making better use of the previous features. Finally, this paper adopts a dual-discriminator generative adversarial network on the network structure, which can force the generated image to retain the intensity of the infrared image and the texture detail information of the visible image at the same time. Substantial qualitative and quantitative experimental analysis of infrared and visible image pairs on three public datasets show that compared with state-of-the-art fusion methods, the proposed MSAt-GAN network has comparable outstanding fusion performance in subjective perception and objective quantitative measurement.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
17  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
20  
ISSN:  
2199-4536  
eISSN:  
2198-6053  
入藏号:  
WOS:000786774500002  
Indexed Date:  
2023-04-07  
Institution Address:  
Kunming Univ Sci & Technol, Fac Informat Engn & Automat, Kunming 650500, Yunnan, Peoples R China  
Institution Address:  
Kunming Univ Sci & Technol, Key Lab Applicat Comp Technol Yunnan Prov, Kunming 650500, Yunnan, Peoples R China  
Institution Address:  
Jiangnan Univ, Sch Artificial Intelligence & Comp Sci, Wuxi 214122, Jiangsu, Peoples R China  
Institution:  
Institution:  
Institution:

### 377. QP-adaptive compressed video super-resolution with coding priors

摘要: Video super-resolution algorithms have found widespread applications as post-processing techniques in down- sampling based coding methods. With the advancements in deep learning techniques, video super-resolution has achieved remarkable success. However, applying existing video super-resolution methods to compressed videos requires training specific models for various quantization parameters (QPs), significantly increasing the resource consumption for model training and compromising their practical utility. To address this issue, we propose a QP-adaptive network for compressed video super-resolution based on coding priors (QPAN). Firstly, we design a QP modulation module (QPMM), which can utilize the frame-wise QP to recalibrate feature maps. Then, on the basis of QPMM, an adaptive multi-scale prior fusion module (Ada-MSPFM) and an adaptive enhancement modulation module (Ada-EMM) are constructed. The former effectively integrates multi-scale features from spatial coding priors in the bitstream and multi-scale features from the decoded video frames. And the latter improves the expressive ability of the network by leveraging QP modulation and reinforcing feature flow adaptively. Extensive experiments demonstrate the highly flexible and adaptive of our proposed method, which exhibits superior reconstruction performance compared to state-of-the-art video super-resolution algorithms.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0165-1684  
eISSN:  
1872-7557  
入藏号:  
WOS:001411374000001  
Indexed Date:  
2025-02-07  
Institution Address:  
Sichuan Univ, Coll Elect & Informat Engn, Chengdu 610065, Peoples R China  
Institution:

### 378. Segmentation of Breast Tubules in H&E Images Based on a DKS-DoubleU-Net

摘要: The formation of breast tubules plays an important role in the pathological grading of breast cancer. Breast tubules surrounded by a large number of epithelial cells are located in the subcutaneous tissue of the chest. The shapes of breast tubules are various, including tubular, round, and oval, which makes the process of breast tubule segmentation a difficult task. Deep learning technology, capable of learning complex data structures via efficient representation, could help pathologists accurately detect breast tubules in hematoxylin and eosin (H&E) stained images. In this paper, we propose a deep learning model named DKS-DoubleU-Net to accurately segment breast tubules with complex appearances in H&E images. The proposed DKS-DoubleU-Net model suggests using a DenseNet module as the encoder of the second subnetwork of DoubleU-Net, which utilizes dense features between layers and strengthens the propagation of features extracted in all previous layers, in order to better discover the intrinsic characteristics of breast tubules with complex structures and diverse shapes. Moreover, a feature fusing module called Kernel Selecting Module (KSM) is inserted before each output layer of the two U-Net branches of the DoubleU-Net, to implement a multiscale feature fusion via a self-adaptive kernel selecting for the sake of accurate segmentation of breast tubules in different sizes. The experiments on the public BRACS dataset and a private clinical dataset have shown that our model achieves better segmentation performance, compared to the state-of-art models of U-Net, DoubleU-Net, ResUnet++, HRNet, and DeepLabV3+. Specifically, on the public BRACS dataset, our method produced an F1-Score of 92.98%, which outperforms the F1-Score of U-Net, DoubleU-Net, and HRNet by 4.24%, 0.37%, and 1.68%, respectively, and is much better than performances of DeepLabV3+ and ResUnet++ by 7.83% and 23.84%, respectively. On the private clinic dataset, the proposed model achieved an F1-Score of 73.13%, which has shown an improvement of 10.31%, 1.89%, 4.88%, 15.47%, and 31.1% to the performances of the U-Net, DoubleU-Net, HRNet, DeepLabV3+, and ResUnet++, respectively. Superior performance could also be observed when comparing the proposed DKS-DoubleU-Net with the others using the metrics of Dice and mIou.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
2314-6133  
eISSN:  
2314-6141  
入藏号:  
WOS:000868908100005  
PubMed ID:  
36246965  
Indexed Date:  
2022-10-26  
Institution Address:  
Shaanxi Normal Univ, Sch Comp Sci, Xian, Peoples R China  
Institution Address:  
Shandong Junteng Med Technol Co LTD, Jinan, Peoples R China  
Institution:  
Institution:  
Shandong Junteng Med Technol Co LTD

### 379. Low-light face detection method based on the cross fusion of high-and

摘要: Face images captured under low-light conditions suffer from significant noise and low contrast,which negatively impact the accuracy of existing face detection systems.In addition,existing low-light image detection algorithms struggle to extract information from small facial areas.To tackle these issues,this paper proposes a two-stage face detection algorithm based on deep learning.This algorithm enhances low-light images before initiating the detection process using an established low-light image enhancement method.The objective is to enhance the ability of the detection method to extract facial information.Thus,a new cross-fusion method of high-and low-frequency channel features is designed.The first step involves using a separable module for high-and low-frequency channel features,enabling the separation of different scale features.These separated features are then cross-fused to improve the ability of the network to extract high-frequency details and low-frequency color information.This,in turn,improves the performance of the detection network.The comparative and ablation experiments validate the effectiveness of the proposed method.The experimental results demonstrate that our method surpasses the baseline method by 4.0% mAP.  
摘要:  
低光条件下捕获的人脸图像存在着噪声严重、对比度低等不足,极大影响了现有人脸检测器的准确性,另外,现有的低光图像检测算法欠缺小区域人脸信息的提取能力。为此,提出一种基于深度学习的两阶段人脸检测算法,即利用现有的低光图像增强算法对人脸图像进行增强后再进行检测。为提升检测算法对人脸信息的提取能力,设计一种新型的高低频通道特征交叉融合方法,该方法首先使用高低频通道特征可分离模块分离出不同尺度特征的高低频信息,然后对上述信息进行交叉融合,提升网络提取高频细节信息和低频色域信息的能力,进而提高检测网络的性能。对比试验和消融试验验证了该研究方法的有效性,试验结果表明该方法优于基准方法4.0% mAP。  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1673-4785  
入藏号:  
CSCD:7718962  
Indexed Date:  
2024-07-26  
Institution Address:  
山西大学大数据科学与产业研究院;;山西大学计算机与信息技术学院, ;;, 太原;;太原, 山西;;山西 030006;;030006, 中国  
Institution Address:  
山西大学大数据科学与产业研究院;;山西大学计算机与信息技术学院;;山西大学, ;;;;计算智能与中文信息处理教育部重点实验室, 太原;;太原;;太原, 山西;;山西;;山西 030006;;030006;;030006, 中国  
Institution Address:  
山西大学大数据科学与产业研究院, 太原, 山西 030006, 中国  
Institution Address:  
Institute of Big Data Science and Industry,Shanxi University;;School of Computer and Information Technology,Shanxi University, ;;, Taiyuan;;Taiyuan, ;; 030006;;030006  
Institution Address:  
Institute of Big Data Science and Industry,Shanxi University;;School of Computer and Information Technology,Shanxi University;;Shanxi University, ;;;;Key Laboratory of Computational Intelligence and Chinese Information Processing, Taiyuan;;Taiyuan;;Taiyuan, ;;;; 030006;;030006;;030006  
Institution Address:  
Institute of Big Data Science and Industry,Shanxi University, Taiyuan, Shanxi 030006, China  
Institution:  
山西大学大数据科学与产业研究院;;山西大学计算机与信息技术学院  
Institution:  
山西大学大数据科学与产业研究院;;山西大学计算机与信息技术学院;;山西大学  
Institution:  
山西大学大数据科学与产业研究院  
Institution:  
Institute of Big Data Science and Industry,Shanxi University;;School of Computer and Information Technology,Shanxi University  
Institution:  
Institute of Big Data Science and Industry,Shanxi University;;School of Computer and Information Technology,Shanxi University;;Shanxi University  
Institution:  
Institute of Big Data Science and Industry,Shanxi University

### 380. MAL-Net: Multiscale Attention Link Network for accurate eye center

摘要: In recent years, due to wide potential applications like human-computer interactions, eye center detection has received growing research interest. Although numerous approaches have been proposed, achieving high accuracy in the wild remains challenging because of the variations in appearance, shape, and illumination. In this paper, we formulate the eye center detection problem as an approximate segmentation task. We propose a novel method to detect pupil positions from a face image using a deep neural network with synthetic data. In particular, we introduce a deep neural network named Multiscale-Attention-Link Network (MAL-Net), where we design a Link Attention Module (LAM) and a novel Multiscale Link Structure (MLS) for accurate and robust eye center detection. Besides, a weighted loss is proposed to make the deep model pay more attention to eye centers during training. Furthermore, to address the problem of insufficient training data with enough variations in eye shape and illumination, we propose a GAN-based method named shape-GAN to generate synthetic eye images with various shapes for training. The proposed MAL-Net is evaluated on widely-used benchmarks such as TFV, GI4E, and BioID. The results demonstrate that our proposed method outperforms state-of-the-art methods for eye center detection.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1077-3142  
eISSN:  
1090-235X  
入藏号:  
WOS:001034308700001  
Indexed Date:  
2023-08-05  
Institution Address:  
Sun Yat Sen Univ, Sch Intelligent Syst Engn, Shenzhen 518107, Peoples R China  
Institution:

### 381. Driver behavior detection via adaptive spatial attention mechanism

摘要: Drivers still play an important role in driving safety despite the presence of driverless vehicles. Over the last few years, millions of deaths are due to traffic accidents, and more than half of these accidents worldwide are caused by distracted driving. Therefore, driver behavior detection during driving is crucial. A novel driver behavior detection system based on the adaptive spatial attention mechanism is proposed in this study. This system realizes the extraction of adaptive discriminative spatial regions of driver images by cascading multiple attentionbased convolution neural networks. Feature representation in each subnetwork is extracted from the output layer, and the discriminative region of the input image is cropped using class activation maps. The obtained region is then fed into the next subnetwork to highlight important region for improving the system performance. The model starts from full images and iteratively crops the region adaptively from coarse to fine to extract the feature representation at multiscales. Finally, the k-nearest neighbor classifier is applied to classify the cascaded multiscale features and obtain the category of driver behavior. The systems are evaluated on a driver behavior recognition database captured in actual driving environments. Experimental results indicate that our systems can achieve superior recognition performance to other state-of-the-art methods and can run in real-time with simplified structure and model in our platform.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
27  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
28  
ISSN:  
1474-0346  
eISSN:  
1873-5320  
入藏号:  
WOS:000663592900008  
Indexed Date:  
2021-07-02  
Institution Address:  
Shandong Jianzhu Univ, Sch Mech & Elect Engn, Jinan 250101, Peoples R China  
Institution Address:  
Shandong Labor Vocat & Tech Coll, Jinan 250300, Peoples R China  
Institution Address:  
Shandong Univ, Joint SDU NTU Ctr Artificial Intelligence Res C F, Jinan 250101, Peoples R China  
Institution Address:  
Nanyang Technol Univ, Sch Mech & Aerosp Engn, Singapore 639798, Singapore  
Institution Address:  
Shandong Univ, Sch Mech Engn, Minist Educ, Key Lab High Efficiency & Clean Mech Mfg, Jinan 250061, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Fac Engn, Hong Kong, Peoples R China  
Institution Address:  
Shandong Promote Mech & Elect Technol Co Ltd, Jinan 250000, Peoples R China  
Institution:  
Institution:  
Shandong Labor Vocat & Tech Coll  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Shandong Promote Mech & Elect Technol Co Ltd

### 382. CGS-Net: A classification-guided framework for automated infection

摘要: Automated segmentation of lung lesions in CT images of COVID-19 based on deep learning holds great potential for comprehending the advancement of the disease and establishing suitable treatment approaches. However, the complex background, indistinct boundaries, varying sizes and distributions of infected regions, and high similarity to other lung diseases pose substantial challenges. To address these issues, we propose a joint deep learning-based framework, named Classification-Guided Segmentation Network (CGS-Net), for COVID-19 segmentation. The framework comprises a classification and a segmentation sub-network, which are trained sequentially. Initially, the classification sub-network learns the feature information of COVID-19 and other pneumonia classifications. Subsequently, the trained classification sub-network assists in the training of the segmentation sub-network. In addition, several key modules are employed to construct the network, including the Multi-scale Feature Mapping Module (MSFM), the Context Information Module (Context), and the Axial Attention Fusion Module (AAFM). The MSFM employs dilated convolutions to extract multi-scale features with an emphasis on spatial and channel information. The Context combines strip pooling and average pooling to aggregate contextual and global information. Finally, the AAFM is employed to incorporate contextual information through an axial attention mechanism. Extensive experimental results demonstrate that our proposed method outperforms competing approaches in the segmentation task.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0899-9457  
eISSN:  
1098-1098  
入藏号:  
WOS:001141987300001  
Indexed Date:  
2024-01-20  
Institution Address:  
Coll Big Data & Informat Engn, Power Syst Engn Res Ctr, Minist Educ, Guiyang, Peoples R China  
Institution Address:  
Guizhou Univ, State Key Lab Publ Big Data, Guiyang, Peoples R China  
Institution Address:  
Coll Big Data & Informat Engn, Power Syst Engn Res Ctr, Minist Educ, Guiyang 550025, Peoples R China  
Institution:  
Coll Big Data & Informat Engn  
Institution:  
Institution:  
Coll Big Data & Informat Engn

### 383. Lightweight object detection model in remote sensing image by combining

摘要: Objective Remote sensing image object detection plays an important role in military security, maritime traffic supervision, intelligent monitoring, and other fields. Remote sensing images are different from natural images. Most remote sensing images are taken at altitudes ranging from several kilometers to tens of thousands of meters. Therefore, the scale of target objects in remote sensing images is large. Most of the target objects are small, such as small vehicles. The other target objects are huge, such as ships. The angles of the objects in the remote sensing images are distributed arbitrarily because of the shooting angle. Therefore, this scenario is a huge challenge for the feature extraction network in remote sensing image target detection, particularly in complex backgrounds. Given the continuous improvement in the computing power of hardware devices and the rapid development of deep learning theory, large and ultralarge object detection networks have been continuously proposed in recent years to improve detection accuracy. Although these detection networks have strong representation learning capabilities, they ignore the cost-effectiveness gained from the relationship of detection accuracy with model calculation amount and the number of parameters. Moreover, real-time detection requirements are difficult to achieve, and the number of parameters and amount of calculation are very limited in model deployment. In addition, most of the general target detection models are designed for natural field datasets. The detection effect in remote sensing image target detection is unsatisfactory, particularly for densely arranged objects. The traditional horizontal box object detection cannot achieve precise detection, such as ships in port and cars in parking lots. Aiming at the above problems, a lightweight rotating box remote sensing image object detection model(YOLO-RMV4)is designed. Method In the experiment, the open-source datasets DOTA2.0, FAIR1M, and HRSC2016 are used as the basic datasets. Moreover, four common vehicles, including a ship, a plane, a small vehicle, and a large vehicle, are selected as objects. A aerial images of vehicle ship and plane(AVSP)dataset is prepared after preprocesses, such as filtering, segmentation, conversion, and relabeling, are performed. This dataset contains 19 406 images of 1 024 \* 1 024 and 637 466 object instances. The AVSP data labels are divided into HBB and OBB(HBB is the horizontal box annotation, and OBB is the rotating box annotation), where OBB is represented by eight parameters. YOLO-RMV4 is improved based on the Mobile- Netv3 network. Adding an efficient channel attention(ECA)mechanism module with excellent performance in the feature extraction network, appropriately expanding the network scale, adding the SPPF module after the feature extraction network, and adding the path aggregation network(PANet)result in multiscale fusion of the extracted features of the backbone network, thereby providing the network with rich and reliable target features. In the network detection head, multiscale detection technology is used to deal with target objects of different sizes. More than half of the objects in the dataset are small targets. Thus, the detection after four times of downsampling is added, resulting in 4, 8, 16, and 32 times of downsampling. Moreover, the small target loss is given a high weight. The smooth circular label is added to the angle prediction in the detection head, which converts the angle regression problem into a classification problem. Thus, the distance between the predicted angle and the real angle can be measured, and the angle periodicity problem is solved. This scenario results in a precise bounding box positioning. Moreover, the anchor size is designed according to the characteristics of the dataset. We use random cropping, flipping, mosaic technique, and other data augmentation approaches in the training.  
摘要:  
目的遥感图像目标检测在国防安全、智能监测等领域扮演着重要的角色。面对遥感图像中排列密集且方向任意分布的目标,传统水平框目标检测不能实现精细定位,大型和超大型的目标检测网络虽然有强大表征学习能力,但是忽略了模型准确率与计算量、参数量之间的性价比,也满足不了实时检测的要求,庞大的参数量和计算量在模型部署上也非常受限,针对以上问题,设计了一种轻量级的旋转框遥感图像目标检测模型(YOLO-RMV4)。方法对原MobileNetv3网络进行改进,在特征提取网络中加入性能更好的通道注意力机制模块(efficient channel attention,ECA),并且对网络规模进行适当扩展,同时加入路径聚合网络(path aggregation network,PANet),对主干网络提取特征进行多尺度融合,为网络提供更丰富可靠的目标特征。网络检测头中则采用多尺度检测技术,来应对不同尺寸的目标物体,检测头中的角度预测加入了环形圆滑标签(circular smooth label,CSL),将角度回归问题转换为分类问题,从而使预测角度和真实角度之间的距离可以衡量。结果将提出的检测模型在制备的AVSP(aerial images of vehicle ship and plane)数据集上进行实验验证,并对主流的7种轻量级网络模型进行了对比实验,相比RYOLOv5l,该模型大小(5.3 MB)仅为RYOLOv5l(45.3 MB)的1/8,平均精度均值(mean average precision,mAP)提高了1.2%,平均召回率(average recall,AR)提高了1.6%。并且mAP和AR均远高于其他的轻量级网络模型。同时也对各个改进模块进行了消融实验,验证了不同模块对模型性能的提升程度。结论本文提出的模型在轻量的网络结构下辅以多尺度融合和旋转框检测,使该模型在极有限参数量下实现实时推理和高精度检测。  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-8961  
入藏号:  
CSCD:7560498  
Indexed Date:  
2023-12-23  
Institution Address:  
大连海事大学航运经济管理学院, 大连, 辽宁 116000, 中国  
Institution Address:  
School of Maritime Economics and Management, Dalian Maritime University, Dalian, Liaoning 116000, China  
Institution:  
大连海事大学航运经济管理学院  
Institution:  
School of Maritime Economics and Management, Dalian Maritime University

### 384. 6D pose estimation based on mask location and hourglass network

摘要: Objective 6D pose estimation is a core problem in 3D object detection and reconstruction. Traditional pose estimation methods usually cannot handle textureless objects. Many post processing procedures have been employed to solve this issue, but they lead to a decline in pose estimation speed. To achieve a fast, single-shot solution, a 6D object pose estimation algorithm based on mask location and heat maps is proposed in this paper. In the prediction of the method, masks are first employed to locate objects, which can reduce the error caused by occlusion. To accelerate mask generation, you only look once v3(YOLOv3)network is used as the backbone. The algorithm presented in this paper does not require any post processing. Our neural network directly predicts the location of key points at a fast speed. Method Our algorithm mainly consists of the following steps. First, a segmentation network structure in object detection is used to generate masks. To speed up this process, YOLOv3 is used as the network backbone. Based on the original detection, a branch structure is added by the segmentation network, and deconvolution is used to extract features under different resolutions. Moreover, 1 \*1, 3 \*3, and 1 \*1 kernel size convolution layers are added to each deconvolution. Finally, these features are fused and used for generating object target and mask map by the mean square error as the loss function in the regression loss. Second, an hourglass network is used to predict key points for each object. A form of encoding and decoding is adopted by the hourglass network. In the encoding stage, down sampling and the residual module are used to reduce the scale and extract features, respectively. Up sampling is used to restore the scale during the decoding. Each level of scale passes through the residual module, and the residual module extracts features without changing the data size. To prevent the feature map from losing local information when the scale is enlarged, a multiscale feature constraint is proposed. Two branches are split to retain the original scale information before each down sampling, and a skip layer containing only one convolution kernel of 1 is used. Stitching is performed at the same scale after one up sampling. Four different resolutions used in convolution are spliced into the up sampling, and the initial feature map is combined with the up sampled feature map. The hourglass network is not directly up sampled to the same resolution size as the network input to obtain the heat map by performing regression. Instead, the hourglass network is used as relay supervision, which restricts the final heat map result from the residual network. Finally, the 6D pose of the object is recovered through the perspective-n-point algorithm. Result In the experimental part, the challenging Linemod datasets are used to evaluate our algorithm. The Linemod dataset has 15 models and is difficult to detect due to the complexity of the object scene. The proposed method is compared with state-of-the-art methods in terms of 3D average distance(ADD)errors and 2D projection error. Results show that the ADD of the paper can reach 82.7%, which is 10% higher than that of the existing heat map method such as Betapose. A 98.9% projection accuracy is reached, and a 4% improvement in 2D projection error is achieved. On symmetrical objects, feature points are selected by Betapose method by considering the symmetry of objects to improve the pose accuracy. As a comparison, feature points are extracted by our algorithm by using the sift method without any symmetry knowledge. However, the results of our algorithm on symmetrical objects are still higher than those of Betapose.  
摘要:  
目的6D姿态估计是3D目标识别及重建中的一个重要问题。由于很多物体表面光滑、无纹理,特征难以提取,导致检测难度大。很多算法依赖后处理过程提高姿态估计精度,导致算法速度降低。针对以上问题,本文提出一种基于热力图的6D物体姿态估计算法。方法首先,采用分割掩码避免遮挡造成的热力图污染导致的特征点预测准确率下降问题。其次,基于漏斗网络架构,无需后处理过程,保证算法具有高效性能。在物体检测阶段,采用一个分割网络结构,使用速度较快的YOLOv3(you only look once v3)作为网络骨架,目的在于预测目标物体掩码分割图,从而减少其他不相关物体通过遮挡带来的影响。为了提高掩码的准确度,增加反卷积层提高特征层的分辨率并对它们进行融合。然后,针对关键点采用漏斗网络进行特征点预测,避免残差网络模块由于局部特征丢失导致的关键点检测准确率下降问题。最后,对检测得到的关键点进行位姿计算,通过PnP(perspective-n-point)算法恢复物体的6D姿态。结果在有挑战的Linemod数据集上进行实验。实验结果表明,本文算法的3D误差准确性为82.7%,与热力图方法相比提高了10%;2D投影准确性为98.9%,比主流算法提高了4%;同时达到了15帧/ s的检测速度。结论本文提出的基于掩码和关键点检测算法不仅有效提高了6D姿态估计准确性,而且可以维持高效的检测速度。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1006-8961  
入藏号:  
CSCD:7220551  
Indexed Date:  
2022-08-05  
Institution Address:  
浙江工业大学计算机科学与技术学院, 杭州, 浙江 310023, 中国  
Institution Address:  
杭州师范大学信息科学与技术学院, 杭州, 浙江 311121, 中国  
Institution Address:  
College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, Zhejiang 310023, China  
Institution Address:  
School of Information Science and Technology, Hangzhou Normal University, Hangzhou, Zhejiang 311121, China  
Institution:  
浙江工业大学计算机科学与技术学院  
Institution:  
杭州师范大学信息科学与技术学院  
Institution:  
College of Computer Science and Technology, Zhejiang University of Technology  
Institution:  
School of Information Science and Technology, Hangzhou Normal University

### 385. Development of Multiscale 3D Residual U-Net to Segment Edematous Adipose

摘要: Data annotation is often a prerequisite for applying deep learning to medical image segmentation. It is a tedious process that requires substantial guidance from experienced physicians. Adipose tissue labeling on CT scans is particularly time-consuming because adipose tissue is present throughout the entire body. One possible solution is to create inaccurate annotations from conventional (non-deep learning) adipose tissue segmentation methods. This work demonstrates the development of a deep learning model directly from these inaccurate annotations. The model is a multi-scale 3D residual U-Net where the encoder path is composed of residual blocks and the decoder path fuses multi-scale feature maps from different layers of decoder blocks. The training set consisted of 101 patients and the testing set consisted of 14 patients. Ten patients with anasarca were purposely added to the testing dataset as a stress test to evaluate model generality. Anasarca is a medical condition that leads to the generalized accumulation of edema within subcutaneous adipose tissue. Edema creates heterogeneity inside the adipose tissue which is absent in the training data. In comparison with a baseline method of manual annotations, the Dice coefficient improved significantly from 73.4 +/- 14.1% to 80.2 +/- 7.1% (p < 0.05). The model trained on inaccurate annotations improved the accuracy of adipose tissue segmentation by 7% without the need for any manual annotation.  
会议名称:  
18th International Symposium on Medical Information Processing and Analysis  
会议日期:  
NOV 09-11, 2022  
会议地点:  
Valparaiso, CHILE  
会议赞助商:  
SIPAIM Fdn; Adv Ctr Elect & Elect Engn; Univ Concepc; Synopsys; Millennium Inst Intelligent Healthcare Engn; Voxel Hlth Care; Univ Tecnica Federico Santa Maria; Pontificia Univ Catolica Chile; Univ Valparaiso; SPIE  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0277-786X  
eISSN:  
1996-756X  
ISBN:  
978-1-5106-6254-4; 978-1-5106-6255-1  
入藏号:  
WOS:001002734500029  
Indexed Date:  
2023-09-07  
Institution Address:  
NIH, Imaging Biomarkers & Comp Aided Diag Lab, Radiol & Imaging Sci Clin Ctr, Bethesda, MD 20892 USA  
Institution:

### 386. Le-SKT: Lightweight traffic density estimation method based on

摘要: The traffic density estimation algorithm usually runs on an embedded platform, and its efficiency is very important to the implementation of applications. However, most of the existing methods rely on heavy backbone networks, require high-performance hardware platform support and take a long time to calculate, which are difficult prerequisites to apply in practice. In consideration of the aforementioned problems, this paper proposes a lightweight traffic density estimation method based on structured knowledge transfer (Le-SKT). This method consists of a teacher network and student network. In the teacher network, in order to solve the problem of variable vehicle scales in video images, a multi scale fusion structure is designed, which can fully fuse the low-level and high-level features of an image. To solve the problem of background interference in video images, a fusion module based on an attention mechanism is constructed to enhance the useful traffic flow feature information and suppress the background information. At the same time, the fused traffic flow feature information is dilated convoluted to expand the receptive field and obtain richer contextual vehicle feature information, which generates a trained teacher network. Finally, a lightweight student network is generated by using the structured knowledge of the teacher network, which solves the problems of long computing times and high hardware requirements. The MAE and GAME of this method are reduced to 3.92 and 5.69, respectively in the TRANCOS dataset and 3.74 and 5.57, respectively in the VisDrone2019 vehicle dataset, respectively; The parameter quantity of the student network is only 0.65% of that of the teacher network, which greatly reduces the scale of the neural network model and is suitable for an embedded computing platform with low performance.(c) 2022 Elsevier Inc. All rights reserved.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0020-0255  
eISSN:  
1872-6291  
入藏号:  
WOS:000817167500014  
Indexed Date:  
2022-07-08  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao 266590, Peoples R China  
Institution Address:  
Shandong Univ Finance & Econ, Dongfang Coll, Tai An 271000, Peoples R China  
Institution Address:  
Nanjing Univ Sci & Technol, Coll Comp Sci & Engn, Nanjing 210096, Peoples R China  
Institution:  
Institution:  
Shandong Univ Finance & Econ  
Institution:

### 387. RepCrack: An efficient pavement crack segmentation method based on

摘要: The detection of pavement cracks is a significant challenge in the field of road maintenance. In response to this challenge, the industry has witnessed a growing trend towards the utilization of deep learning-based methods for crack detection. However, existing approaches utilize complex network structures which, while improving accuracy, compromise detection efficiency and hinder practical application. This paper proposes a novel pixel-level pavement crack segmentation model, RepCrack, which addresses the current limitations in accuracy and efficiency in pavement crack detection tasks. The proposed model is based on large kernel convolution and multiscale feature extraction, in order to extract more detailed pavement crack information during the training phase. The model building process incorporates structural re-parameterization techniques, including the conversion of multi-branch structures to single on-path structures and the shortening of long sequence structures, which effectively reduces the complexity of the model. Furthermore, to enhance the model's capacity to extract features, an improved coordinate attention mechanism is incorporated to augment the network's focus on the crack region. To validate the detection effect, publicly available datasets, including the Crack Forest Dataset (CFD), German Asphalt Pavement Distress (GAPs) dataset, and Crack Tree dataset, were employed. The experimental results demonstrate that our proposed method exhibits a notable superiority over existing crack detection techniques. For instance, RepCrack, with a parameter count of 5.8 million, attains an Intersection over Union (IoU) of 63.8% on the CFD dataset and attains a state-of-the-art methodology (SOTA) performance, exhibiting a 5% enhancement in accuracy and a 68.7% reduction in inference latency in comparison to SegNext.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001377934200001  
Indexed Date:  
2024-12-21  
Institution Address:  
Soochow Univ, Sch Rail Transportat, Suzhou, Jiangsu, Peoples R China  
Institution Address:  
Intelligent Urban Rail Engn Res Ctr, Yangzhou, Jiangsu, Peoples R China  
Institution:  
Institution:  
Intelligent Urban Rail Engn Res Ctr

### 388. Multi-scale feature fusion for prediction of IDH1 mutations in glioma

摘要: Background and objective: Mutations in isocitrate dehydrogenase 1 (IDH1) play a crucial role in the prognosis, diagnosis, and treatment of gliomas. However, current methods for determining its mutation status, such as immunohistochemistry and gene sequencing, are difficult to implement widely in routine clinical diagnosis. Recent studies have shown that using deep learning methods based on pathological images of glioma can predict the mutation status of the IDH1 gene. However, our research focuses on utilizing multi-scale information in pathological images to improve the accuracy of predicting IDH1 gene mutations, thereby providing an accurate and cost-effective prediction method for routine clinical diagnosis. Methods: In this paper, we propose a multi-scale fusion gene identification network (MultiGeneNet). The network first uses two feature extractors to obtain feature maps at different scale images, and then by employing a bilinear pooling layer based on Hadamard product to realize the fusion of multi-scale features. Through fully exploiting the complementarity among features at different scales, we are able to obtain a more comprehensive and rich representation of multi-scale features. Results: Based on the Hematoxylin and Eosin stained pathological section dataset of 296 patients, our method achieved an accuracy of 83.575 % and an AUC of 0.886, thus significantly outperforming other single-scale methods. Conclusions: Our method can be deployed in medical aid systems at very low cost, serving as a diagnostic or prognostic tool for glioma patients in medically underserved areas.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0169-2607  
eISSN:  
1872-7565  
入藏号:  
WOS:001222967100001  
PubMed ID:  
38518408  
Indexed Date:  
2024-05-23  
Institution Address:  
Chinese Acad Sci, Shenzhen Inst Adv Technol, Shenzhen 518055, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Shenzhen Coll Adv Technol, Shenzhen 518055, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Guangdong Prov Clin Res Ctr Canc, Dept Pathol,Canc Ctr, State Key Lab Oncol South China, Guangzhou 510060, Peoples R China  
Institution Address:  
Sorbonne Univ, INRIA, AP HP, Paris Brain Inst ICM,Inserm,CNRS, F-75013 Paris, France  
Institution:  
Institution:  
Institution:  
Institution:

### 389. A multiscale enhanced pavement crack segmentation network coupling

摘要: Road pavement cracks are a critical factor affecting the health conditions of road pavements. Accurate crack detection contributes to providing data support for road maintenance measures. Compared to conventional crack detection algorithms, deep learning based crack segmentation methods have practical significance for road maintenance and traffic safety management due to their high precision and automation. However, existing deep learning methods often suffer from segmentation accuracy loss due to the varying crack sizes and the presence of stains on the pavement with spatial characteristics similar to cracks, leading to the misclassification of cracks. Therefore, this study proposed a multiscale enhanced road pavement crack segmentation network (MS-CrackSeg) by coupling spectral and spatial information to detect pavement cracks from unmanned aerial vehicle (UAV) hyperspectral imagery. MS-CrackSeg can simultaneously learn the spatial and rich spectral features of cracks in the hyperspectral imagery, improving the discrimination of those targets with similar spatial features to cracks compared to previous approaches. Moreover, the Multiscale Self-Attention-like Feature Extraction Module (MSSA) is introduced to extract and fuse multiscale crack features to enhance the crack detection. Experiments on a dataset consisting of 1031 hyperspectral images demonstrated the superior crack segmentation of the proposed method compared to the comparative methods. In particular, the proposed method achieved the highest F1-score of 0.74 and mean Intersection over Union of 0.79, indicating an exceptional performance. The developed approach offers improved data support for road maintenance measures and has validated the advantages of UAV hyperspectral imagery in road crack segmentation. The annotated hyperspectral dataset and the code for MS-CrackSeg network are available at https://github.com/williamchen-x/MS-CrackSeg.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1569-8432  
eISSN:  
1872-826X  
入藏号:  
WOS:001219348800001  
Indexed Date:  
2024-05-17  
Institution Address:  
Peking Univ, Inst Remote Sensing & Geog Informat Syst, 5 Summer Palace Rd, Beijing 100871, Peoples R China  
Institution Address:  
Minist Educ China, Engn Res Ctr Earth Observat & Nav CEON, 5 Summer Palace Rd, Beijing, Peoples R China  
Institution:  
Institution:  
Minist Educ China

## Other / Unspecified

### 1. A Recognition Model Based on Multiscale Feature Fusion for Needle-Shaped

摘要: To solve the problem that traditional seed recognition methods are not completely suitable for needle-shaped seeds, such as Bidens L., in agricultural production, this paper proposes a model construction idea that combines the advantages of deep residual models in extracting high-level abstract features with multiscale feature extraction fusion, taking into account the depth and width of the network. Based on this, a multiscale feature fusion deep residual network (MSFF-ResNet) is proposed, and image segmentation is performed before classification. The image segmentation is performed by a popular semantic segmentation method, U2Net, which accurately separates seeds from the background. The multiscale feature fusion network is a deep residual model based on a residual network of 34 layers (ResNet34), and it contains a multiscale feature fusion module and an attention mechanism. The multiscale feature fusion module is designed to extract features of different scales of needle-shaped seeds, while the attention mechanism is used to improve the ability to select features of our model so that the model can pay more attention to the key features. The results show that the average accuracy and average F1-score of the multiscale feature fusion deep residual network on the test set are 93.81% and 94.44%, respectively, and the numbers of floating-point operations per second (FLOPs) and parameters are 5.95 G and 6.15 M, respectively. Compared to other deep residual networks, the multiscale feature fusion deep residual network achieves the highest classification accuracy. Therefore, the network proposed in this paper can classify needle-shaped seeds efficiently and provide a reference for seed recognition in agriculture.

### 2. Yolov4 High-Speed Train Wheelset Tread Defect Detection System Based on

摘要: The Yolov4 detection algorithm does not sufficiently extract local semantic and location information. This study aims to solve this problem by proposing a Yolov4-based multiscale feature fusion detection system for high-speed train wheel tread defects. First, multiscale feature maps are obtained from a feature extraction backbone network. The proposed multiscale feature fusion network then fuses the underlying features of the original three scales. These fused features contain more defect semantic information and location details. Based on the fused features, a path aggregation network is used to fuse feature maps at different resolutions, with an improved loss function that speeds up the convergence of the network. Experimental results show that the proposed method is effective at detecting defects in the wheel treads of high-speed trains.

### 3. RAFF-Net: An improved tongue segmentation algorithm based on residual

摘要: Objective Due to the complexity of face images, tongue segmentation is susceptible to interference from uneven tongue texture, lips and face, resulting in traditional methods failing to segment the tongue accurately. To address this problem, RAFF-Net, an automatic tongue region segmentation network based on residual attention network and multiscale feature fusion, was proposed. It aims to improve tongue segmentation accuracy and achieve end-to-end automated segmentation. Methods Based on the UNet backbone network, different numbers of ResBlocks combined with the Squeeze-and-Excitation (SE) block was used as an encoder to extract image layered features. The decoder structure of UNet was simplified and the number of parameters of the network model was reduced. Meanwhile, the multiscale feature fusion module was designed to optimize the network parameters by combining a custom loss function instead of the common cross-entropy loss function to further improve the detection accuracy. Results The RAFF-Net network structure achieved Mean Intersection over Union (MIoU) and F1-score of 97.85% and 97.73%, respectively, which improved 0.56% and 0.46%, respectively, compared with the original UNet; ablation experiments demonstrated that the improved algorithm could contribute to the enhancement of tongue segmentation effect. Conclusion This study combined the residual attention network with multiscale feature fusion to effectively improve the segmentation accuracy of the tongue region, and optimized the input and output of the UNet network using different numbers of ResBlocks, SE block, multiscale feature fusion and weighted loss function, increased the stability of the network and improved the overall effect of the network.

### 4. Chinese named entity recognition method based on multiscale feature

摘要: In response to the problems of low recognition accuracy and poor recognition efficiency in traditional methods, the paper proposes a Chinese named entity recognition method based on multiscale feature fusion. Firstly, the similarity between each word is calculated using the literal similarity algorithm to obtain synonyms of Chinese named entities. Then, the Chinese named entity features are obtained, including character features, character shape features, binary character features, and word similarity features, through multiscale feature fusion to obtain the Chinese named entity feature set. Finally, the target Chinese named entity for recognition is obtained by matching vocabulary, compressing vocabulary vectors, and integrating character vectors, and the CRF is used to implement Chinese named entity recognition. The experimental results show that the recognition time of this method is only 4.0 s, with a precision rate of up to 99.9% and a recall rate of up to 99.2%.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1755-8301  
eISSN:  
1755-831X  
入藏号:  
WOS:001245423900010  
Indexed Date:  
2024-06-19  
Institution Address:  
Yongcheng Vocat Coll, Dept Culture & Arts, Yongcheng 476600, Peoples R China  
Institution:  
Yongcheng Vocat Coll

### 5. An integrated deep multiscale feature fusion network for aeroengine

摘要: Most RUL prediction methods can only extract single-scale features, ignoring significant details at other scales and layers. These methods are all constructed using one type of model, and do not use the advantages of different models. An integrated deep multiscale feature fusion network (IDMFFN) for aeroengine RUL prediction using multisensor data is proposed in this study. Two-dimensional samples are constructed using multisensor data with multiple time cycles. Multiscale feature extraction blocks are designed to learn different-scale features using convolutional filters of different sizes. A multiscale feature concatenated block is constructed to integrate multiscale features from different layers. A GRUbased high-level feature fusion block is built to replace the traditional fully connected layer, and can leverage powerful temporal feature learning for feature fusion. A novel activation function Mish is used to construct the network. A simulated turbofan engine dataset was used to verify the effectiveness of the network. The results suggest that the IDMFFN can predict RUL more accurately than existing methods. (C) 2021 Published by Elsevier B.V.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
70  
在中国科学引文数据库中的被引频次:  
3  
被引频次合计:  
73  
ISSN:  
0950-7051  
eISSN:  
1872-7409  
入藏号:  
WOS:000718126500010  
Indexed Date:  
2021-11-24  
Institution Address:  
Northwestern Polytech Univ, Sch Civil Aviat, Xian 710072, Peoples R China  
Institution Address:  
Harbin Inst Technol, Sch Astronaut, Harbin 150001, Peoples R China  
Institution Address:  
AECC Sichuan Gas Turbine Estab, Mianyang 621010, Sichuan, Peoples R China  
Institution:  
Institution:  
Institution:  
AECC Sichuan Gas Turbine Estab

### 6. RDBL-Net: detection of foreign objects on transmission lines based on

摘要: Timely detection of foreign objects on transmission lines is the key to ensuring transmission lines' safe and stable operation. This paper takes the dataset provided by the Guangzhou Pazhou Algorithm Competition as the data basis, proposing a transmission line foreign object detection model. First, a residual connection module combined with deformable convolution is introduced into the backbone, aiming to improve the feature extraction capability of the model for transmission line foreign objects. Second, a multiscale feature fusion structure is designed to enhance the fusion effect of multiscale features. Subsequently, a learnable position encoding is introduced into the multiscale feature interaction module to enhance the model's ability to cope with complex environments' interference. Finally, the effectiveness of the proposed method is demonstrated through experiments.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1748-1279  
eISSN:  
1748-1287  
入藏号:  
WOS:001426610100005  
Indexed Date:  
2025-02-26  
Institution Address:  
Northeast Elect Power Univ, Sch Comp Sci, Jilin 132012, Jilin, Peoples R China  
Institution Address:  
Northeast Elect Power Univ, Jilin 132012, Jilin, Peoples R China  
Institution Address:  
Liaoning Univ Technol, Sch Econ & Management, Jinzhou 121001, Liaoning, Peoples R China  
Institution Address:  
China Logist Co Ltd, Harbin Branch, Harbin 100073, Heilongjiang, Peoples R China  
Institution Address:  
Jilin Investment Promot Ctr, Informat Sect, Jilin 132011, Jilin, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
China Logist Co Ltd  
Institution:  
Jilin Investment Promot Ctr

### 7. Adaptive multiscale feature for object detection

摘要: In object detection, multiscale features are necessary to deal with objects with different size. Using Feature Pyramid Network (FPN) as the backbone network is a very popular paradigm in existing object detectors, we call this paradigm FPN+. However, feature fusion of FPN is insufficient to express object of similar size but different appearance due to the unidirectional feature fusion. We motivate and present Adaptive Multiscale Feature (AMF), a new multiscale feature fusion method with bidirectional feature fusion, using to solve the one-direction fusion of FPN. AMF module fuses multiscale features from two aspects: (1) shattering features by the way of CLSM; (2) fusing features by the way of channel-wise attention. The proposed AMF improves the expression ability of multiscale features of the backbone network, and effectively improves the performance of the object detector. The proposed feature fusion method can be added to all object detector, such as the one-stage detector, the two-stage detector, anchor-based detector and anchor-free based detector. Experimental results on the COCO 2014 dataset show that the proposed AMF module performs the popular FPN based detector. Whether anchored-free based detectors or anchored based detectors, performance of detector can be improved through AMF. And the resulting best model can achieve a very competitive mAP on COCO 2014 dataset.  
(c) 2021 Elsevier B.V. All rights reserved.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
在 BIOSIS Citation Index 中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
被引频次合计:  
13  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000652818400013  
Indexed Date:  
2021-06-05  
Institution Address:  
Chinese Acad Sci, Xian Inst Opt & Precis Mech, Key Lab Spectral Imaging Technol CAS, Xian 710119, Shaanxi, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution Address:  
Chinese Acad Sci, Xian Inst Opt & Precis Mech XIOPM, Key Lab Ultrafast Photoelect Diagnost Technol, Xian 710119, Shaanxi, Peoples R China  
Institution:  
Institution:  
Institution:

### 8. CSASNet-A Crop Leaf Disease Identification Method Based on Improved

摘要: In identifying crop leaf diseases, Due to the complex nature of the disease symptoms. There may be variations in disease symptoms with similar characteristics and similarities in disease symptoms with different elements. This can make it challenging to differentiate between various diseases. CSASNet is a hybrid classification model proposed in this paper that combines the attention and multiscale feature fusion mechanisms. The model first incorporates the multiscale feature fusion module atrous spatial pyramid pooling (ASPP) into the ShuffleNetV2 network structure. This enriches the network with disease-specific multiscale feature information. Additionally, the model combines the special group-wise enhance (SGE) attention mechanism module to enhance the weight of disease spot feature information. Lastly, the leaky ReLU function replaces the original ReLU activation function. This allows the model to reduce damaging feature loss during training. The paper presents a design of multiple cross-validation experiments for comparison. The experimental results suggest that the improved model was used for disease leaf identification and showed an accuracy improvement on different crops. Compared to Convnext and MobileNetV2, the CSASNet model demonstrates higher recognition accuracy.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0146-4116  
eISSN:  
1558-108X  
入藏号:  
WOS:001300711800008  
Indexed Date:  
2024-09-04  
Institution Address:  
Northeast Elect Power Univ, Sch Comp Engn, Jilin 132012, Peoples R China  
Institution:

### 9. DMFF-Net: A dual encoding multiscale feature fusion network for ovarian

摘要: Ovarian cancer is a serious threat to the female reproductive system. Precise segmentation of the tumor area helps the doctors to further diagnose the disease. Automatic segmentation techniques for abstracting high-quality features from images through autonomous learning of model have become a hot research topic nowadays. However, the existing methods still have the problem of poor segmentation of ovarian tumor details. To cope with this problem, a dual encoding based multiscale feature fusion network (DMFF-Net) is proposed for ovarian tumor segmentation. Firstly, a dual encoding method is proposed to extract diverse features. These two encoding paths are composed of residual blocks and single dense aggregation blocks, respectively. Secondly, a multiscale feature fusion block is proposed to generate more advanced features. This block constructs feature fusion between two encoding paths to alleviate the feature loss during deep extraction and further increase the information content of the features. Finally, coordinate attention is added to the decoding stage after the feature concatenation, which enables the decoding stage to capture the valid information accurately. The test results show that the proposed method outperforms existing medical image segmentation algorithms for segmenting lesion details. Moreover, the proposed method also performs well in two other segmentation tasks.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2296-2565  
入藏号:  
WOS:000917827400001  
PubMed ID:  
36711337  
Indexed Date:  
2023-02-10  
Institution Address:  
Tiangong Univ, Sch Life Sci, Tianjin, Peoples R China  
Institution Address:  
Tiangong Univ, Sch Control Sci & Engn, Tianjin, Peoples R China  
Institution Address:  
China Univ Petr, Coll Comp Sci & Technol, Qingdao, Peoples R China  
Institution Address:  
Qingdao Univ, Dept Gynecol, Affiliated Hosp, Qingdao, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 10. AM-Net: A Network With Attention and Multiscale Feature Fusion for Skin

摘要: Accurate segmentation of skin lesions is crucial for the diagnosis and treatment of skin diseases. Common problems in dermoscopic medical images, such as inconsistent scale information, blurred small-size boundaries, and irregular shapes of lesion areas, limit the performance of existing methods. To this end, we proposed AM-Net to effectively alleviate the above problems. To address the issue of inconsistent scale information, we designed the multiscale feature integration module (MFIM) and the multiscale feature fusion module (MFFM) as the basic modules for network encoding and decoding. The MFIM integrates the feature information of different scales to enhance feature extraction, and the MFFM processes the multiscale information in parallel to effectively fuse the image features. To address the problem of blurred small-size boundaries, we designed the detail boundary enhancement attention module (DBEAM), which strengthens key details and boundary information in images using an attention-mechanism-weighted approach. To address the problem of irregular shapes of skin lesions, we designed the spatial-channel feature fusion module (SCFFM) to effectively combine feature information at different levels in the encoder-decoder for interaction, enhancing the segmentation capability of irregularly shaped lesion areas. The experiments on the ISIC-2016, ISIC-2017, and PH2 datasets demonstrate that our method achieves Dice coefficients of 0.9329, 0.873, and 0.9149, respectively, outperforming existing advanced methods and effectively achieving precise segmentation of dermoscopic image lesions. Our code is available at https://github.com/8yike/AM-Net.git.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001438192900007  
Indexed Date:  
2025-03-14  
Institution Address:  
Guangxi Univ Sci & Technol, Sch Automat, Liuzhou 545006, Peoples R China  
Institution Address:  
Guangxi Univ Sci & Technol, Dept Med, Liuzhou 545006, Peoples R China  
Institution:  
Institution:

### 11. Pyramid Vision Transformer based on Bidirectional Multiscale Feature

摘要: Pests damage crops, thereby resulting in reduced production and quality. In order to reduce the damage of pests, it is important to timely implement pest control according to various types of pests. Existing pest classification is identified by the naked eye of an experienced farmer or expert. This requires considerable time and cost, so an unmanned pest classification technology is needed. In this paper, a new bidirectional multiscale feature fusion method for pest image classification is proposed. In particular, in order to supplement the scale vulnerability of the existing PVT model, the multi-scale features extracted from the PVT backbone are applied bidirectionally by four fusion methods. Through the experimental results, the fusion method with the highest performance is derived and it is shown that the proposed bidirectional multiscale feature fusion method can improve the pest image classification performance.  
摘要:  
해충은 작물을 가해하여 생산량 감소와 품질 저하를 수반한다. 따라서 해충 피해를 최소화하기 위해서는 해충 종류에 따른 방제 작업을 적기에 실행하는 것이 중요하다. 기존의 해충 분류는 경험이 풍부한 농부나 전문가의 육안에 의해 감별되었다. 이는 많은 시간과 비용을 요구하므로 무인 기반의 해충 분류 기술이 필요하다. 따라서 이 논문에서는 해충 영상 분류를 위한, 양방향 멀티스케일 특징 퓨전 방식을 새롭게 제안하고자 한다. 특히, 기존의 PVT 모델의 스케일 취약성 문제를 보완하고자, PVT 백본에서 추출된 멀티스케일 특징을 양방향으로 4가지의 퓨전 방식을 적용하고자 한다. 실험 결과를 통해, 가장 성능이 높은 퓨전 방식을 도출하고 제안한 양방향 멀티스케일 특징 퓨전 방식이 해충 영상 분류 성능을 개선할 수 있음을 선보이고자 한다.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1598-8619  
入藏号:  
KJD:ART002891104  
Indexed Date:  
2023-01-25

### 12. MFFN: An Underwater Sensing Scene Image Enhancement Method Based on

摘要: Vision-guided autonomous underwater vehicles based on remote sensing play an important role in ocean missions. However, some problems exist in underwater visual perception, such as color distortion, low contrast, and fuzzy details, which restrict the applications of underwater visual tasks. Most of the state-of-the-art image enhancement methods are still limited in scene adaptability, recovery accuracy, and real-time processing. To solve these problems, we propose an underwater sensing scene image enhancement method called a multiscale feature fusion network (MFFN). To extract the multiscale feature, the measure merging the feature extraction module, the feature fusion module, and the attention reconstruction module is designed. This measure can also enhance the adaptability and visual effect of the scene. Moreover, we propose multiple objective functions for supervised training to match the nonlinear mapping. Based on the qualitative and quantitative evaluations, the proposed method produces competitive performance compared with some state-of-the-art methods, and the perception and statistical quality of underwater images are enhanced effectively.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
39  
在中国科学引文数据库中的被引频次:  
4  
在 BIOSIS Citation Index 中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
42  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000766298800025  
Indexed Date:  
2022-03-25  
Institution Address:  
Macau Univ Sci & Technol, Fac Informat Technol, Macau, Peoples R China  
Institution Address:  
Jinan Univ, Modern Educ Technol Ctr, Zhuhai 519070, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Sch Atmospher Sci, Planetary Environm & Astrobiol Res Lab, Zhuhai 519088, Peoples R China  
Institution:  
Institution:  
Institution:

### 13. Joint margin adaption and multiscale feature fusion for COVID-19 CT

摘要: Segmenting COVID-19 from CT images remains a challenging task due to the structural characteristics, including complex and diverse infected regions, high inter-class similarity, and intra-class variability, which make it difficult to capture detailed semantic information such as edges and textures effectively. The present study introduces a novel methodology for segmenting COVID-19 CT images, consisting of three distinct stages. The proposed approach is primarily based on the margin adaptive deep supervised feature fusion network (MADFNet) architecture, which adopts an encoder-decoder framework. In this strategy, the segmentation outcomes obtained in the previous phase are utilized as supervised terms to facilitate the semantic segmentation process in the subsequent stage. In MADFNet, firstly, the channel edge detection and spatial edge detection modules capture the rich semantic information of edge context from different dimensions to obtain the fine texture information and semantic dependencies, respectively. Then, the encoder feature fusion module is designed to aggregate multiscale feature representations at the same level, facilitating information integration across different channels and enabling adaptive detection of structural features such as edges and textures. Finally, the decoder feature fusion module integrates multiscale feature dependencies from different levels to alleviate the semantic discrepancies of features at different levels and accurately locate lesion regions. The single-class segmentation Dice scores of MADFNet on the three experimental datasets are 0.831, 0.825, and 0.821, respectively, and the results of single-class as well as multi-class lesion segmentation on COVID-19 CT images validated the feasibility of our work.  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:001154162100001  
Indexed Date:  
2024-02-09  
Institution Address:  
NanChang HangKong Univ, Sch Software, Nanchang 330063, Peoples R China  
Institution Address:  
Univ Tehran, Coll Engn, Sch Surveying & Geospatial Engn, Tehran, Iran  
Institution Address:  
Wenzhou Univ, Key Lab Intelligent Informat Safety & Emergency Zh, Wenzhou 325035, Peoples R China  
Institution Address:  
Jiangxi Canc Hosp, Dept Med Imaging, Nanchang 330029, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Jiangxi Canc Hosp

### 14. SSGDD-YOLO: Multiscale Feature Fusion and Multiattention-Based YOLO for

摘要: Surface defect detection is essential for ensuring the product quality of smartphone screen glass. In this work, a smartphone screen glass defect detection model based on an enhanced YOLOv7 framework with multiscale feature fusion and multiattention, named SSGDD-you only look once (YOLO) is proposed. In the developed SSGDD-YOLO model, the branch fusion block (BFB) is integrated low-level features from multiple scales through parallel processing, to enhance the details in lower level features for minimizing the information loss as less as possible. Furthermore, the SPPCSPC module of the head is improved as the SPPCSPC-I module, by replacing the standard max pooling with local importance-based pooling (LIP) that reflects the importance of features. The developed SPPCSPC-I module allows the network to automatically learn adaptive importance weights of features during downsampling, enhancing the multiscale feature extraction capability with diverse receptive fields. Finally, a contour-mixed attention block (C-MAB) is inserted into the feature fusion section of the network, which enhances spatial and channel information of features to reduce target information loss, improving the representation capability. Experiments are conducted using a challenging real-world defect image dataset gathered from a smartphone screen glass inspection line in an industrial plant. Results show the proposed SSGDD-YOLO model can achieve the highest mAP of 62.46% among all compared methods.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001422065500020  
Indexed Date:  
2025-02-26  
Institution Address:  
Zhejiang Sci Tech Univ, Sch Informat Sci & Engn, Hangzhou 310018, Peoples R China  
Institution:

### 15. A hybrid multiscale feature fusion model for enhanced cardiovascular

摘要: Cardiovascular arrhythmia, characterized by irregular heart rhythms, can lead to severe complications such as stroke and heart failure if not detected promptly. Traditional arrhythmia classification methods often struggle with class imbalance and fail to capture critical multiscale temporal and spatial features, leading to suboptimal diagnostic performance. To address these challenges, this study proposes a Hybrid Multiscale Feature Fusion (HMFF) strategy that combines ConvNeXt-X models with advanced data balancing techniques. Specifically, we explore two fusion approaches, Hybrid Feature Fusion (HFF) and HMFF, leveraging ConvNeXtTiny, ConvNeXtSmall, and ConvNeXtBase architectures. To mitigate class imbalance, we integrate SMOTE-Tomek Link (STL) and Random Oversampling (RO) strategies, ensuring robust representation of minority classes. Using the MIT-BIHA Dataset, our HMFF approach demonstrated substantial improvements over baseline models, which achieved an accuracy of 97.35% without feature fusion or data balancing. The HFF and HMFF models improved accuracies to 98.44% and 98.49%, respectively, even without balancing techniques. When combined with RO, accuracies increased to 99.11% and 99.25%, and further improved with STL to 99.15% and 99.30%. These results highlight the significant advancements provided by the HMFF strategy, offering an effective and scalable solution for enhancing arrhythmia detection and diagnosis in clinical settings.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2590-1230  
入藏号:  
WOS:001422238000001  
Indexed Date:  
2025-02-21  
Institution Address:  
Int Univ Business Agr & Technol, Dept Comp Sci & Engn, Dhaka, Bangladesh  
Institution:

### 16. Generative Adversarial Network With Dual Multiscale Feature Fusion for

摘要: The performance of intelligent fault diagnosis models heavily depends on the amount of monitoring data available. In the situations of monitoring data insufficient for fault diagnosis, generative adversarial networks (GANs) can augment the existing data to supplement data scarcity, which is a promising approach to improve diagnostic accuracy. However, the quality of the generated samples greatly affects the effectiveness of this method. To address this issue, this article proposes a dual multiscale feature fusion (MSFF) GAN to ensure the similarity between generated and real samples and also to improve the diversity of the generated samples. Specifically, a multiscale feature extraction and fusion module is designed to integrate multiscale feature extraction and fusion. A multiscale feature decision fusion module is constructed to avoid the loss of decision-sensitive features in different healthy states. The design of the dual MSFF enhances the learning ability of the generation model and guarantees the similarity between the generated and real samples. A reconstruction network is established to restrain the error of the latent vectors reconstructed by the generated samples, thereby preventing the overfitting of the generated samples and improving their diversity. Experimental results demonstrate that the proposed model has advantages in the similarity, diversity, and effectiveness of the generated samples, significantly improving the performance of intelligent fault diagnosis.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
7  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001081486200022  
Indexed Date:  
2023-11-12  
Institution Address:  
Xi An Jiao Tong Univ, Key Lab Educ Minist Modern Design & Rotor Bearing, Xian 710049, Peoples R China  
Institution Address:  
Univ Technol Sydney, Sch Mech & Mechatron Engn, Ultimo, NSW 2007, Australia  
Institution Address:  
Univ British Columbia, Sch Engn, Kelowna, BC V1V 1V7, Canada  
Institution:  
Institution:  
Institution:

### 17. Self-Attention Guidance and Multiscale Feature Fusion-Based UAV Image

摘要: Object detection on unmanned aerial vehicle (UAV) images is a recent research hotspot. Existing object detection methods have achieved good results on general scenes, but there are inherent challenges with UAV images. The detection accuracy of UAV images is limited by complex backgrounds, significant scale differences, and densely arranged small objects. To solve these problems, we propose a UAV image object detection network based on self-attention guidance and multiscale feature fusion (SGMFNet). First, we design a global-local feature guidance (GLFG) module. This module can effectively combine local information and global information, which makes the model focus on the object area and reduces the impact of complex background. Second, an improved parallel sampling feature fusion (PSFF) module is designed to efficiently fuse multiscale features. Third, we design an inverse-residual feature enhancement (IFE) module, which is embedded in the front of the newly added detection head to enhance feature extraction on small objects. Finally, we conduct a large number of experiments on the VisDrone2019 dataset. The results show that the proposed SGMFNet outperforms other popular methods and has achieved good results in many scenarios.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
40  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
40  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000979491800003  
Indexed Date:  
2023-05-29  
Institution Address:  
Shijiazhuang Tiedao Univ, Shijiazhuang 050043, Peoples R China  
Institution:

### 18. Small-Target Traffic Sign Detection Based on Multiscale Feature Fusion

摘要: A small-target traffic sign detection algorithm based on multiscale feature fusion is proposed to address the limited effectiveness of the existing target detection algorithms in detecting traffic signs with small sizes or inconspicuous features. First, a bidirectional adaptive-feature pyramid network is designed to extract all detail features and jump connections to enhance multiscale feature fusion. Second, a dual-head detection structure is proposed for the scale characteristics of small targets, focusing on small-target feature information while reducing the number of model parameters. Next, using the Wise-IoU v3 bounding box loss function and a dynamic nonmonotonic focusing mechanism, the harmful gradients generated by low-quality examples are reduced by employing the anchor-box gradient gain allocation strategy. Finally, coordinate convolution (CoordConv) is incorporated into the feature extraction network to enhance the spatial awareness of the model by improving the network's focus on coordinate information. The experimental results on the Tsinghua-Tencent 100K dataset show that the improved model has a mean average precision (mAP) of 87.7%, which is a 2. 2 percentage points improvement over YOLOv5s. Moreover, the number of parameters is only 6. 3x10(7), thereby achieving a detection effect with fewer parameters and higher accuracy.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001284198700037  
Indexed Date:  
2024-09-18  
Institution Address:  
North China Univ Sci & Technol, Coll Elect Engn, Tangshan 063210, Hebei, Peoples R China  
Institution Address:  
Tianjin Chengjian Univ, Sch Control & Mech Engn, Tianjin 300384, Peoples R China  
Institution:  
Institution:

### 19. Small-Target Traffic Sign Detection Based on Multiscale Feature Fusion

摘要: A small-target traffic sign detection algorithm based on multiscale feature fusion is proposed to address the limited effectiveness of the existing target detection algorithms in detecting traffic signs with small sizes or inconspicuous features. First, a bidirectional adaptive-feature pyramid network is designed to extract all detail features and jump connections to enhance multiscale feature fusion. Second, a dual-head detection structure is proposed for the scale characteristics of small targets, focusing on small-target feature information while reducing the number of model parameters. Next, using the Wise-IoU v3 bounding box loss function and a dynamic nonmonotonic focusing mechanism, the harmful gradients generated by low-quality examples are reduced by employing the anchor-box gradient gain allocation strategy. Finally, coordinate convolution (CoordConv) is incorporated into the feature extraction network to enhance the spatial awareness of the model by improving the network's focus on coordinate information. The experimental results on the Tsinghua-Tencent 100K dataset show that the improved model has a mean average precision (mAP) of 87.7%, which is a 2.2 percentage points improvement over YOLOv5s. Moreover, the number of parameters is only 6.3\*10~7, thereby achieving a detection effect with fewer parameters and higher accuracy.  
摘要:  
针对现有目标检测算法对小尺寸或特征不明显的交通标志检测效果较差的问题,提出一种基于多尺度特征融合的小目标交通标志检测算法。首先,设计一种双向自适应特征金字塔网络,充分利用细节特征和跳跃连接,增强多尺度特征融合;其次,针对小目标的尺度特点提出双头检测结构,聚焦小目标特征信息,同时减少模型的参数量;再次,使用Wise-IoU v3边界框损失函数,结合动态非单调聚焦机制,利用锚框梯度增益分配策略,减小低质量示例产生的有害梯度;最后,在特征提取网络中融入坐标卷积(CoordConv),提升网络对坐标信息的关注程度,从而增强模型的空间感知能力。在Tsinghua-Tencent 100K数据集上的实验结果表明,改进后的模型的平均精度均值(mAP)为87.7%,较YOLOv5s提升了2.2百分点,且参数量仅为6.3\*10~7,达到了参数量更少、精度更高的检测效果。  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7783058  
Indexed Date:  
2024-10-18  
Institution Address:  
华北理工大学电气工程学院, 唐山, 河北 063210, 中国  
Institution Address:  
天津城建大学控制与机械工程学院, 天津 300384, 中国  
Institution Address:  
College of Electrical Engineering,North China University of Science and Technology, Tangshan, Hebei 063210, China  
Institution Address:  
School of Control and Mechanical Engineering,Tianjin Chengjian University, Tianjin 300384, China  
Institution:  
华北理工大学电气工程学院  
Institution:  
天津城建大学控制与机械工程学院  
Institution:  
College of Electrical Engineering,North China University of Science and Technology  
Institution:  
School of Control and Mechanical Engineering,Tianjin Chengjian University

### 20. Enhanced Image Retrieval Using Multiscale Deep Feature Fusion in

摘要: In recent years, deep-network-based hashing has gained prominence in image retrieval for its ability to generate compact and efficient binary representations. However, most existing methods predominantly focus on high-level semantic features extracted from the final layers of networks, often neglecting structural details that are crucial for capturing spatial relationships within images. Achieving a balance between preserving structural information and maximizing retrieval accuracy is the key to effective image hashing and retrieval. To address this challenge, we introduce Multiscale Deep Feature Fusion for Supervised Hashing (MDFF-SH), a novel approach that integrates multiscale feature fusion into the hashing process. The hallmark of MDFF-SH lies in its ability to combine low-level structural features with high-level semantic context, synthesizing robust and compact hash codes. By leveraging multiscale features from multiple convolutional layers, MDFF-SH ensures the preservation of fine-grained image details while maintaining global semantic integrity, achieving a harmonious balance that enhances retrieval precision and recall. Our approach demonstrated a superior performance on benchmark datasets, achieving significant gains in the Mean Average Precision (MAP) compared with the state-of-the-art methods: 9.5% on CIFAR-10, 5% on NUS-WIDE, and 11.5% on MS-COCO. These results highlight the effectiveness of MDFF-SH in bridging structural and semantic information, setting a new standard for high-precision image retrieval through multiscale feature fusion.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2313-433X  
入藏号:  
WOS:001404227300001  
PubMed ID:  
39852333  
Indexed Date:  
2025-01-30  
Institution Address:  
High Sch Comp Sci, Sidi Bel Abbes 22000, Algeria  
Institution Address:  
Sidi Bel Abbes Univ, Telecommun Dept, RCAM Lab, Sidi Bel Abbes 22000, Algeria  
Institution:  
Institution:

### 21. Many heads are better than one: A multiscale neural information feature

摘要: The neural information at different scales exhibits spatial representations and the corresponding features are believed to be conducive for neural encoding. However, existing neural decoding studies on multiscale feature fusion have rarely been investigated. In this study, a multiscale neural information feature fusion framework is presented and we integrate these features to decode spatial routes from multichannel recordings. We design a goal-directed spatial cognitive experiment in which the pigeons need to perform a route selection task. Multichannel neural activities including spike and local field potential (LFP) recordings in the hippocampus are recorded and analyzed. The multiscale neural information features including spike firing rate features, LFP time frequency energy features, and functional network connectivity features are extracted for spatial route decoding. Finally, we fuse the multiscale feature to solve the neural decoding problem and the results indicate that feature fusion operation improves the decoding performance significantly. Ten-fold cross-validation result analysis shows a promising improvement in the decoding performance using fusing multiscale features by an average of 0.04-0.11 at least than using any individual feature set alone. The proposed framework investigates the possibility of route decoding based on multiscale features, providing an effective way to solve the neural information decoding problems.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0361-9230  
eISSN:  
1873-2747  
入藏号:  
WOS:000792761000001  
PubMed ID:  
35293319  
Indexed Date:  
2022-05-18  
Institution Address:  
Zhengzhou Univ, Sch Elect Engn, Henan Key Lab Brain Sci & Brain Comp Interface Te, Zhengzhou, Henan, Peoples R China  
Institution Address:  
Ruhr Univ Bochum, Fac Psychol, Inst Cognit Neurosci, Dept Biopsychol, Bochum, Germany  
Institution:  
Institution:

### 22. Multiscale Feature Fusion Approach for Dual-Modal Object Detection

摘要: Object detection based on visible images is difficult to adapt to complex lighting conditions such as low light, no light, strong light, etc., while object detection based on infrared images is greatly affected by background noise. Infrared objects lack color information and have weak texture features, which pose a greater challenge. To address these problems, a dual-modal object detection approach that can effectively fuse the features of visible and infrared dual-modal images is proposed. A multiscale feature attention module is proposed, which can extract the multiscale features of the input IR and RGB images separately. Meanwhile, channel attention and spatial pixel attention is introduced to focus the multiscale feature information of dual-modal images from both channel and pixel dimensions. Finally, a dual-modal feature fusion module is proposed to adaptively fuse the feature information of dual-modal images. On the large-scale dual-modal image dataset DroneVehicle, compared with the benchmark algorithm YOLOv5s using visible or infrared single-modal image detection, the proposed algorithm improves the detection accuracy by 13.42 and 2.27 percentage points, and the detection speed reaches 164 frame/s, with ultra-real-time end-to-end detection capability. The proposed algorithm effectively improves the robustness and accuracy of object detection in complex scenes, which has good application prospects.  
摘要:  
基于可见光图像的目标检测,难以适应弱光、无光、强光等复杂光照条件,而基于红外图像的目标检测,受背景噪声影响大,且红外目标缺乏颜色信息,纹理细节特征弱,给目标检测带来较大挑战。对此,提出了一种能够有效融合可见光与红外图像特征的双模态目标检测方法。对输入的成对的双模态图像分别提取其初级特征;提出了多尺度特征注意力模块,对输入的红外与可见光图像分别提取其多尺度局部特征,并引入通道注意力和空间像素注意力,从通道和像素两个维度聚焦双模态图像的多尺度特征信息;提出双模态特征融合模块,对双模态特征信息进行自适应融合,得到双模态图像的多尺度融合特征。在大规模双模态图像数据集DroneVehicle上,与基准算法YOLOv5s利用可见光或红外单模态图像进行检测相比,所提算法检测精度分别提升了13.42和2.27个百分点,同时检测速度达到164 frame/s,具备端到端的实时检测能力。所提算法有效提高了复杂场景下目标检测的鲁棒性和准确性,具有良好的应用前景。  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7804362  
Indexed Date:  
2024-11-23  
Institution Address:  
陆军工程大学指挥控制工程学院, 南京, 江苏 210007, 中国  
Institution Address:  
College of Command and Control Engineering,Army Engineering University of PLA, Nanjing, Jiangsu 210007, China  
Institution:  
陆军工程大学指挥控制工程学院  
Institution:  
College of Command and Control Engineering,Army Engineering University of PLA

### 23. Attention-enhanced multiscale fusion for face mask-wearing detection

摘要: Mask wearing is known as one of the most effective and convenient methods for preventing the transmission of viruses. Accurately determining whether a person is wearing a mask correctly is crucial in curbing the widespread transmission of viruses in public places. To address the limitations of traditional approaches, such as insufficient feature extraction and limited generalizability, an attention-enhanced multiscale fusion approach is proposed for assisting face mask-wearing detection. We first incorporate the simple, parameter-free attention module, adaptively spatial feature fusion module, and wise intersection over union loss function into a multiscale feature fusion framework, constructing the YSAW mask-wearing detection network, resulting in the development of the YSAW mask-wearing detection network. The network was then trained and evaluated on the WIDER face dataset and MAFA dataset, which includes faces, masks, and face masks. Experimental results demonstrate that the proposed YSAW network achieves mean average precision at intersection over union threshold 0.5 (mAP50=94.3%), indicating superior detection performance in comparison to existing methods.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:001173268100025  
Indexed Date:  
2024-03-11  
Institution Address:  
Jimei Univ, Chengyi Coll, Chinasoft Int Internet Sch, Xiamen, Peoples R China  
Institution Address:  
Jimei Univ, Sch Comp Engn, Xiamen, Peoples R China  
Institution Address:  
Jimei Univ, Sch Ocean Informat Engn, Xiamen, Peoples R China  
Institution:  
Institution:  
Institution:

### 24. Dynamic Multiscale Feature Fusion Method for Underwater Target

摘要: The feature information of small-scale targets is seriously missing under the interference of complex underwater terrain and light refraction. Moreover, the unbalanced distribution of underwater target samples can also affect the accuracy of spatial semantic feature extraction. Aiming at the above problems, this paper proposes a dynamic multiscale feature fusion method for underwater target recognition. Firstly, this paper uses multiscale info noise contrastive estimation (MS-InfoNCE) loss to extract the significant features of the target at 4 scales. Secondly, the method learns the spatial semantic features of the target through a dynamic conditional probability matrix. Finally, this paper designs different feature fusion mechanisms for different scale targets, dynamically fusing multiscale significant features and spatial semantic features to recognize underwater weak targets. The experimental results show that the recognition accuracy of the proposed algorithm is 1.38% higher than that of the existing algorithm when recognizing underwater distorted targets.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1687-725X  
eISSN:  
1687-7268  
入藏号:  
WOS:000835151500001  
Indexed Date:  
2022-08-11  
Institution Address:  
Henan Inst Sci & Technol, Sch Artificial Intelligence, Xinxiang 453003, Peoples R China  
Institution Address:  
Henan Inst Sci & Technol, Sch Informat Engn, Xinxiang 453003, Peoples R China  
Institution:  
Institution:

### 25. Speech emotion recognition based on multimodal and multiscale feature

摘要: Conventional feature extraction methods for speech emotion recognition often suffer from unidimensionality and inadequacy in capturing the full range of emotional cues, limiting their effectiveness. To address these challenges, this paper introduces a novel network model named Multi-Modal Speech Emotion Recognition Network (MMSERNet). This model leverages the power of multimodal and multiscale feature fusion to significantly enhance the accuracy of speech emotion recognition. MMSERNet is composed of three specialized sub-networks, each dedicated to the extraction of distinct feature types: cepstral coefficients, spectrogram features, and textual features. It integrates audio features derived from Mel-frequency cepstral coefficients and Mel spectrograms with textual features obtained from word vectors, thereby creating a rich, comprehensive representation of emotional content. The fusion of these diverse feature sets facilitates a robust multimodal approach to emotion recognition. Extensive empirical evaluations of the MMSERNet model on benchmark datasets such as IEMOCAP and MELD demonstrate not only significant improvements in recognition accuracy but also an efficient use of model parameters, ensuring scalability and practical applicability.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001389232600002  
Indexed Date:  
2025-01-09  
Institution Address:  
Changchun Univ Technol, Coll Comp Sci & Engn, Changchun, Peoples R China  
Institution Address:  
Changchun Normal Univ, Coll Comp Sci & Technol, Changchun, Peoples R China  
Institution:  
Institution:

### 26. Multiscale Feature Fusion Attention Lightweight Facial Expression

摘要: Facial expression recognition based on residual networks is important for technologies related to space human-robot interaction and collaboration but suffers from low accuracy and slow computation in complex network structures. To solve these problems, this paper proposes a multiscale feature fusion attention lightweight wide residual network. The network first uses an improved random erasing method to preprocess facial expression images, which improves the generalizability of the model. The use of a modified depthwise separable convolution in the feature extraction network reduces the computational effort associated with the network parameters and enhances the characterization of the extracted features through a channel shuffle operation. Then, an improved bottleneck block is used to reduce the dimensionality of the upper layer network feature map to further reduce the number of network parameters while enhancing the network feature extraction capability. Finally, an optimized multiscale feature lightweight attention mechanism module is embedded to further improve the feature extractability of the network for human facial expressions. The experimental results show that the accuracy of the model is 73.21%, 98.72%, and 95.21% on FER2013, CK+ and JAFFE, respectively, with a covariance of 10.14 M. Compared with other networks, the model proposed in this paper has faster computing speed and better accuracy at the same time.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
3  
ISSN:  
1687-5966  
eISSN:  
1687-5974  
入藏号:  
WOS:000871485200001  
Indexed Date:  
2022-10-29  
Institution Address:  
Chongqing Univ Technol, Coll Comp Sci & Engn, Chongqing 400054, Peoples R China  
Institution Address:  
Northeastern Univ, Sydney Smart Technol Coll, Shenyang 066004, Peoples R China  
Institution:  
Institution:

### 27. DFLM-YOLO: A Lightweight YOLO Model with Multiscale Feature Fusion

摘要: Object detection algorithms for open water aerial images present challenges such as small object size, unsatisfactory detection accuracy, numerous network parameters, and enormous computational demands. Current detection algorithms struggle to meet the accuracy and speed requirements while being deployable on small mobile devices. This paper proposes DFLM-YOLO, a lightweight small-object detection network based on the YOLOv8 algorithm with multiscale feature fusion. Firstly, to solve the class imbalance problem of the SeaDroneSee dataset, we propose a data augmentation algorithm called Small Object Multiplication (SOM). SOM enhances dataset balance by increasing the number of objects in specific categories, thereby improving model accuracy and generalization capabilities. Secondly, we optimize the backbone network structure by implementing Depthwise Separable Convolution (DSConv) and the newly designed FasterBlock-CGLU-C2f (FC-C2f), which reduces the model's parameters and inference time. Finally, we design the Lightweight Multiscale Feature Fusion Network (LMFN) to address the challenges of multiscale variations by gradually fusing the four feature layers extracted from the backbone network in three stages. In addition, LMFN incorporates the Dilated Re-param Block structure to increase the effective receptive field and improve the model's classification ability and detection accuracy. The experimental results on the SeaDroneSee dataset indicate that DFLM-YOLO improves the mean average precision (mAP) by 12.4% compared to the original YOLOv8s, while reducing parameters by 67.2%. This achievement provides a new solution for Unmanned Aerial Vehicles (UAVs) to conduct object detection missions in open water efficiently.  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2504-446X  
入藏号:  
WOS:001305033400001  
Indexed Date:  
2024-09-09  
Institution Address:  
Donghua Univ, Minist Educ, Coll Informat Sci & Technol, Engn Res Ctr Digitized Text & Fash Technol, Shanghai 201620, Peoples R China  
Institution:

### 28. Retinal vessel segmentation by using AFNet

摘要: Retinal vessel segmentation can obtain rich ocular information which is important for the diagnosis of fundus diseases. To address the problems of existing segmentation methods such as poor capillary segmentation and incorrect segmentation of pathological information, an AFNet vessel segmentation network combining location attention, semantic aggregation module and multi-scale feature fusion module is proposed. Add positional attention to the feature codec block of the network for modeling global dependencies and reducing intra-class inconsistencies, the multiscale feature fusion module is used in the last layer of the coding part to extract multiscale feature information to solve the difficult problem of large variation of retinal vessel width and size, and the designed semantic aggregation module can fully utilize the contextual semantic information to improve the segmentation accuracy of capillaries. Extensive experiments are conducted on three publicly available fundus image databases DRIVE, STARE and CHASE\_DB1, and the results show that AFNet can effectively improve the accuracy of retinal vessel segmentation and achieve better comprehensive performance compared with other methods.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
9  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:000792074600001  
Indexed Date:  
2022-05-18  
Institution Address:  
Guangzhou Univ, Sch Elect & Commun Engn, Guangzhou 511400, Peoples R China  
Institution Address:  
Guangzhou Univ, Sch Mech & Elect Engn, Guangzhou 511400, Peoples R China  
Institution:  
Institution:

### 29. Remote Sensing Image-Matching Network Based on Multiscale Feature Fusion

摘要: Remote sensing image matching is one of the fundamental challenges in earth observation.The complexity and diversity of surface information in remote sensing images often pose difficulties for image matching.To overcome these difficulties,a remote sensing image-matching network based on multiscale feature fusion and importance ranking loss is proposed.This network comprises two parts:a key-point detection network and a feature descriptor extraction network.The key-point detection network has a multilayer convolutional structure based on feature pyramids.This structure is designed to achieve multiscale feature fusion at different network levels.Multiple convolution kernels are used to gradually expand receptive fields at the same level,thereby fully capturing multiscale information in remote sensing images.Furthermore,CBAM is used to aggregate the response graph of the key-point detection network to detect key points with significant scores.The key-point detection network is optimized using the score loss and image block loss,and the feature descriptor sub-extraction network is optimized using the descriptor subloss.The score-importance sorting loss function,descriptor sub-importance sorting loss function,and neighbor mask-based descriptor subloss function are specially designed to ensure that the key points,descriptors,and image blocks used for remote sensing image matching have high repeatability and distinguishability,which improves the accuracy of remote sensing image matching.In this study,many remote sensing images were collected and a remote sensing image-matching dataset was constructed via homography transformation.This dataset was used to experimentally verify the performance of the proposed network model.Compared with traditional image-matching methods or other end-to-end deeplearning image-matching methods,the proposed network model has considerable advantages in remote sensing image matching.  
摘要:  
遥感图像匹配是对地观测的基本问题之一,由于遥感图像中地表信息复杂、尺度多样,往往会对遥感图像匹配造成困难。为此,提出基于多尺度特征融合与重要性排序损失的遥感图像匹配网络。该网络由关键点检测网络和特征描述子提取网络两部分构成。在关键点检测网络中,设计了基于特征金字塔的多层卷积结构,使多尺度特征融合在不同网络层级上实现,并在同一层级中利用多个卷积核逐渐扩大感受野,从而更充分地捕获遥感图像中的多尺度信息。同时,利用CBAM对关键点检测网络的响应图进行聚合,以检测出具有显著得分的关键点。使用分数损失和图像块损失对关键点检测网络进行优化,使用描述子损失对特征描述子提取网络进行优化,并专门设计了分数重要性排序损失函数、描述子重要性排序损失函数,以及基于邻居掩码的描述子损失函数,以保证用于遥感图像匹配的关键点、描述子、图像块具有较高的可重复性、可区分性,从而提高遥感图像匹配的准确性。收集大量遥感图像,通过单应性变换构建遥感图像匹配数据集,并利用该数据集对所提网络模型的性能进行实验验证,发现相比传统图像匹配方法或是其他端到端的深度学习图像匹配方法,所提网络模型在遥感图像匹配中均具有明显优势。  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7841305  
Indexed Date:  
2025-01-19  
Institution Address:  
河北工业大学人工智能与数据科学学院, 天津 300401, 中国  
Institution Address:  
School of Artificial Intelligence,Hebei University of Technology, Tianjin 300401, China  
Institution:  
河北工业大学人工智能与数据科学学院  
Institution:  
School of Artificial Intelligence,Hebei University of Technology

### 30. GLF-Net: A Target Detection Method Based on Global and Local Multiscale

摘要: Due to the influence of weather and shooting height, remote sensing aircraft image has the characteristics of small target samples, fuzzy target samples, and complex background information. Because of these characteristics, most target detection methods applied to original images have the defect of low precision and recall due to insufficient feature extraction. To solve these problems, based on the target features of remote sensing aircraft images, a remote sensing aircraft detection method based on global and local multiscale feature fusion is proposed. The method applies the encoder-decoder architecture, which extracts the local features of the target through the network encoder part, and extracts the global features of the target through the decoder part. The fused features are input into the classifier for evaluation, and a group with high scores is selected as the final detection output. The experimental results show that the proposed method has higher precision, recall, and F1 score. Compared with other detection methods, it has smaller memory usage and fewer parameters.  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000750321500003  
Indexed Date:  
2022-02-10  
Institution Address:  
Harbin Engn Univ, Sch Informat & Commun Engn, Harbin 150001, Peoples R China  
Institution Address:  
Harbin Engn Univ, Minist Ind & Informat Technol, Sch Informat & Commun Engn, Harbin 150001, Peoples R China  
Institution:  
Institution:

### 31. Mask-Guided Attention for Subcategory-Level Sewer Pipe Crack

摘要: In the field of sewage pipeline inspection, the Pipeline Assessment Certification Program (PACP) requires encoding pipe cracks at the subcategory level in practical applications. This study seeks to address the need to improve the accuracy of the crack pixel-level segmentation and subcategory classification problem based on sewer pipe inspection video images. The proposed framework consists of a feature extraction module, a multiscale feature fusion module, a mask-guided attention module, and a subcategory classification module. First, the multiscale feature fusion mechanism is used to integrate the convolution features of different scales to obtain the convolution features rich in both semantic and detailed information, which can improve the characterization ability of cracks and achieve accurate crack detection. Then, mask-guided attention is used to refine features and eliminate background noise, enhancing the feature representation for crack pixels and reducing the ambiguity of the fine-grained crack classification task as well. Finally, the subcategory classification result is obtained based on the feature maps with erased background clutter. The experimental results show that the proposed method achieved F1-scores of 97%, 94%, and 95% for longitudinal, circumferential, and multiple cracks, respectively. The main contribution of the study is the formulation of a multitask collaborative learning framework, which combines pixel-level segmentation and global image-level classification annotation to achieve a desirable performance of sewer crack detection and classification at a subcategory level.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1949-1190  
eISSN:  
1949-1204  
入藏号:  
WOS:001126316600010  
Indexed Date:  
2024-03-21  
Institution Address:  
Jiangsu Univ Sci & Technol, Sch Comp Sci & Engn, Zhenjiang 212003, Jiangsu, Peoples R China  
Institution Address:  
Jiangsu Univ, Sch Elect & Informat Engn, Zhenjiang 212013, Jiangsu, Peoples R China  
Institution Address:  
Oklahoma State Univ, Sch Civil & Environm Engn, Stillwater, OK 74078 USA  
Institution:  
Institution:  
Institution:

### 32. Multiscale and Multidimensional Weighted Network for Salient Object

摘要: Recent advancements have significantly benefited in the area of salient object detection in optical remote sensing images (ORSI-SOD). Given the varying spatial resolutions and complex scenes characteristic of optical remote sensing images (ORSIs), leveraging and integrating features across scales is vital. However, excessive feature integration can introduce significant noise and result in inaccurate saliency mapping. To address this issue, we propose the multiscale and multidimensional weighted network for ORSI-SOD (WeightNet). The network adopts a two-stage design where the first stage generates multiscale weighted information, and the second stage conducts indirect multiscale feature weighted fusion. This design skillfully avoids the lack of scale adaptation and noise interference that may arise from direct multiscale feature fusion. Furthermore, to enhance feature fusion and target localization precision, we introduce the multiscale weighted feature aggregation module (MWFAM) and the multidimensional feature guidance module (MDFGM). MWFAM facilitates multiscale feature fusion while minimizing noise from cross-layer interactions. MDFGM specializes in precise target localization and enhancement of detail and edge information. Additionally, the introduction of a multiscale parallel decoder (MPD) significantly boosts the decoder's capability in identifying targets across various scales. Extensive qualitative and quantitative evaluations on three public ORSI datasets demonstrate the effectiveness and superiority of WeightNet over contemporary state-of-the-art models.  
Web of Science 核心合集中的 "被引频次":  
10  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001237717100011  
Indexed Date:  
2024-06-14  
Institution Address:  
Xi An Jiao Tong Univ, Sch Software Engn, Xian 100091, Peoples R China  
Institution:

### 33. MFFNet: a building change detection method based on fusion of spectral

摘要: Accurate detection and extraction of changes in buildings heights is important in monitoring construction (both legal and illegal) and assessing disasters. It is also important information for updating real 3D scenes. However, when using remote sensing images, shadows, vegetation and objects with similar spectral and morphological characteristics as buildings can cause false detections, omissions and incomplete patch edges. To address this issue, we develop the multiscale feature fusion network for dual-modal data (MFFNet), which has two main aspects: (1) The multi-dual-modal feature fusion module detects changes in features with similar spectral and morphological characteristics as buildings. This mitigates false detections by making the model more aware of areas where the elevation has changed over time. (2) Because building extraction is affected by shadows and vegetation, we designed a multiscale feature shuffle module. It takes multiscale features and establishes relationships between neighbouring pixels using the pixel-shuffle algorithm, then fuses and reorganizes the multiscale features to highlight the relationships between global contexts, thereby mitigating the problem of building occlusion by shadows. Comparative experiments show that MFFNet achieves better results on GF7-CD and 3DCD datasets than other similar methods. The proposed method can more accurately monitor building changes over large areas.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1010-6049  
eISSN:  
1752-0762  
入藏号:  
WOS:001185061400001  
Indexed Date:  
2024-03-20  
Institution Address:  
Chongqing Jiaotong Univ, Sch Smart City, Chongqing, Peoples R China  
Institution Address:  
Minist Nat Resources, Key Lab Monitoring Evaluat & Early Warning Terr Sp, Chongqing, Peoples R China  
Institution Address:  
Minist Nat Resources, Technol Innovat Ctr Spatio temporal Informat & Equ, Chongqing, Peoples R China  
Institution Address:  
Geol Bur Ningxia Hui Autonomous Reg, Yinchuan, Ningxia, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Geol Bur Ningxia Hui Autonomous Reg

### 34. A Small-Sized Object Detection Oriented Multi-Scale Feature Fusion

摘要: Object detection is a well-known task in the field of computer vision, especially the small target detection problem that has aroused great academic attention. In order to improve the detection performance of small objects, in this article, a novel enhanced multiscale feature fusion method is proposed, namely, the atrous spatial pyramid pooling-balanced-feature pyramid network (ABFPN). In particular, the atrous convolution operators with different dilation rates are employed to make full use of context information, where the skip connection is applied to achieve sufficient feature fusions. In addition, there is a balanced module to integrate and enhance features at different levels. The performance of the proposed ABFPN is evaluated on three public benchmark datasets, and experimental results demonstrate that it is a reliable and efficient feature fusion method. Furthermore, in order to validate the applicational potential in small objects, the developed ABFPN is utilized to detect surface tiny defects of the printed circuit board (PCB), which acts as the neck part of an improved PCB defect detection (IPDD) framework. While designing the IPDD, several powerful strategies are also employed to further improve the overall performance, which is evaluated via extensive ablation studies. Experiments on a public PCB defect detection database have demonstrated the superiority of the designed IPDD framework against the other seven state-of-the-art methods, which further validates the practicality of the proposed ABFPN.  
在中国科学引文数据库中的被引频次:  
17  
在 BIOSIS Citation Index 中的被引频次:  
13  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
330  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
343  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000767839800005  
Indexed Date:  
2022-03-24  
Institution Address:  
Xiamen Univ, Dept Instrumental & Elect Engn, Xiamen 361005, Fujian, Peoples R China  
Institution Address:  
Brunel Univ London, Dept Comp Sci, Uxbridge UB8 3PH, Middx, England  
Institution:  
Institution:

### 35. Ancient mural segmentation based on multiscale feature fusion and dual

摘要: To address the fuzzy segmentation boundaries, missing details, small target losses and low efficiency of traditional segmentation methods in ancient mural image segmentation scenarios, this paper proposes a mural segmentation model based on multiscale feature fusion and a dual attention-augmented segmentation model (MFAM). The model uses the MobileViT network, which integrates a coordinate attention mechanism, as the feature extraction backbone network. It attains global and local expression capabilities through self-attention, class convolution, and coordinate attention and focuses on location information to expand the receptive field and achieve improved feature extraction efficiency. An A\_R\_ASPP feature enhancement module is proposed for the attention-optimized residual atrous spatial pyramid pooling module. The module uses residual connections to solve the small target loss problem in murals caused by the excessive sampling rate of atrous convolution and uses a feature attention mechanism to adaptively adjust the feature map weight according to the channel importance levels. A dual attention-enhanced feature fusion module is proposed for multiscale decoder feature fusion to improve the mural segmentation effect. This module uses a cross-level aggregation strategy and an attention mechanism to weight the importance of different feature levels to obtain multilevel semantic feature representations. The model improves the mean intersection over union (MIoU) by 3.06% and the MPA by 1.81% on a mural dataset compared with other models. The model is proven to be effective at improving the segmentation details, efficiency and small target segmentation results produced for mural images, and a new method is proposed for segmenting ancient mural images.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
2050-7445  
入藏号:  
WOS:001163115200002  
Indexed Date:  
2024-02-28  
Institution Address:  
Taiyuan Univ Sci & Technol, Sch Comp Sci & Technol, Taiyuan 030024, Peoples R China  
Institution Address:  
Xinzhou Normal Univ, Dept Comp Sci & Technol, 1 Dunqi East St, Xinzhou 034000, Peoples R China  
Institution Address:  
SEGi Univ, Informat Technol, Petaling Jaya, Selangor, Malaysia  
Institution:  
Institution:  
Institution:

### 36. Small sample remote sensing image segmentation based on multiscale

摘要: Aiming at the problems of high cost of drawing remote sensing image labels and low detection accuracy of remote sensing image under the condition of limited training samples in actual scenes,the deep pyramid attention network (DPA-Net) combining multiscale feature fusion and attention mechanism was proposed to extract buildings and roads from small sample remote sensing images by integrating transfer learning method.Due to limited training samples containing information is limited,first of all,on the basis of DeeplabV3+ network architecture,two sources of low-level features are added to make full use of spatial information of low-level features,and the attention mechanism is used to obtain rich context information and enhance the learning ability of the target channel,reduce the response ability to other targets and noises,and to improve the poor detection effect of the model on small samples.Finally,the transfer learning method using open remote sensing image dataset and small sample dataset for joint training reduces the impact of too few training samples on network learning performance.Experimental results show that the accuracy of the proposed method is improved by 3.69% and the annotation cost can be saved by 1/2.  
摘要:  
针对绘制遥感图像标签成本高、在实际场景中训练样本有限情况下遥感图像检测精度低的问题,提出了一种结合多尺度特征融合与注意力机制的深度金字塔注意力网络(DPA-Net)并集成迁移学习方法进行小样本遥感图像中建筑物与道路的提取.因为有限训练样本包含的信息有限,所以首先在DeeplabV3+网络架构基础上,增加两路低层特征的来源以充分利用低级特征的空间信息,并且利用注意力机制获取丰富的上下信息并增强模型对目标通道的学习能力,降低对其他目标和噪音的响应能力,改善模型在小样本上检测效果差的问题.最后利用公开遥感图像数据集和小样本数据集进行联合训练的迁移学习方法降低训练样本过少对网络学习性能的影响.实验结果表明:本文方法的精度提高了3.69%,可节省1/2的标注成本.  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1671-4512  
入藏号:  
CSCD:7183657  
Indexed Date:  
2022-06-17  
Institution Address:  
中国海洋大学工程学院, 青岛, 山东 266100, 中国  
Institution Address:  
School of Engineering,Ocean University of China, Qingdao, Shandong 266100, China  
Institution:  
中国海洋大学工程学院  
Institution:  
School of Engineering,Ocean University of China

### 37. Egc-yolo: strip steel surface defect detection method based on edge

摘要: Due to imperfect manufacturing crafts and external factors, steel often produces surface defects during manufacturing, seriously influencing its lifespan and availability. It is therefore crucial that surface defects are detected in industrial production. Nevertheless, conventional detection techniques are vulnerable to background interference and feature scale variations when employed to identify defects on strip surfaces. Therefore, we propose an EGC-YOLO model based on YOLOv8 for steel surface defect detection. First, an edge detail enhancement module (EDEM), based on Sobel convolution (SobelConv), is designed and embedded into C2f to capture defective edges and texture better. Second, the generalized dynamic feature pyramid network (GDFPN) is introduced in the neck structure to enhance the multiscale feature fusion. This enables the model to adapt to defects of different sizes and shapes. Finally, the content-guided attention fusion (CGA Fusion) module is employed to optimize the fusion of shallow and deep features for more detection precision. The extensive experimental results illustrate that the accuracy of EGC-YOLO reaches 80.2% mAP on NEU-DET and improves by 3.7% over YOLOv8. The model's inference speed reached 136.4 frames per second (FPS). EGC-YOLO outperforms other models in accuracy and speed for detecting steel surface defects, showcasing its industrial application potential.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001427674400001  
Indexed Date:  
2025-02-26  
Institution Address:  
Xian Univ Sci & Technol, Sch Commun & Informat Engn, Xian 710054, Peoples R China  
Institution Address:  
Xian Key Lab Network Convergence Commun, Xian 710054, Peoples R China  
Institution:  
Institution:  
Xian Key Lab Network Convergence Commun

### 38. SSD with multi-scale feature fusion and attention mechanism

摘要: In the field of the Internet of Things, image acquisition equipment is the very important equipment, which will generate lots of invalid data during real-time monitoring. Analyzing the data collected directly from the terminal by edge calculation, we can remove invalid frames and improve the accuracy of system detection. SSD algorithm has a relatively light and fast detection speed. However, SSD algorithm do not take full advantage of both shallow and deep information of data. So a multiscale feature fusion attention mechanism structure based on SSD algorithm has been proposed in this paper, which combines multiscale feature fusion and attention mechanism. The adjacent feature layers for each detection layer are fused to improve the feature information expression ability. Then, the attention mechanism is added to increase the attention of the feature map channels. The results of the experiments show that the detection accuracy of the optimized model is improved, and the reliability of edge calculation has been improved.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
6  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
2045-2322  
入藏号:  
WOS:001251525900028  
PubMed ID:  
38049437  
Indexed Date:  
2024-07-26  
Institution Address:  
Hunan Univ Technol, Coll Comp & Sci, Zhuzhou, Hunan, Peoples R China  
Institution Address:  
Hunan Univ Technol, Intelligent Informat Percept & Proc Technol, Hunan Prov Key Lab, Zhuzhou, Hunan, Peoples R China  
Institution:  
Institution:

### 39. F-UNet plus plus : Remote Sensing Image Fusion Based on Multipurpose

摘要: The fusion of multispectral (MS) and panchromatic (PAN) images is of great significance for the construction of high-resolution remote sensing images. Because of differences in sensors, no single MS or PAN image can express the complete information of a scene. Therefore, it is a key issue to fuse MS images containing rich spectral content and PAN images with spatial information to construct a high-resolution MS image. In this work, an adaptive shuffle attention (ASA) module and an optimized UNet++ are combined in a fusion-UNet++ (F-UNet++) framework for the problem of MS and PAN image fusion. This ASA module can focus on important information in the mixed domain and adjust the dimensions of tensors. F-UNet++ includes a multiscale feature extraction module, multiscale feature fusion module, and image reconstruction module. The multiscale feature extraction module obtains spectral and spatial information, the multiscale feature fusion module fuses spectral and spatial information, and a composite multi-input image reconstruction module (CMI-UNet++) reconstructs the final image. By combining the ASA attention module, the loss of feature information can be reduced to enhance the fidelity of the spectral and spatial information of the fused image. Experiments show that F-UNet++ is qualitatively and quantitatively superior to current image fusion methods. (The code is available at https://github.com/Josephing/F-UNet).  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000922147200088  
Indexed Date:  
2023-05-22  
Institution Address:  
Yunnan Univ, Engn Res Ctr Cyberspace & Sch Software, Kunming 650091, Yunnan, Peoples R China  
Institution:

### 40. Octave-YOLO: Direct Multi-scale Feature Fusion for Object Detection

摘要: In object detection research, multiscale feature fusioncombining feature maps of different scales to detect objects of varying sizeshas become a critical focus. Network structures like Feature Pyramid Networks (FPNs) and Path Aggregation Networks (PANets) have been developed to address this challenge. PANet, an enhancement of FPN, integrates both top-down and bottom-up pathways, leading to significant improvements in object detection performance. However, during multiscale feature fusion, PANet’s upscaling and downscaling processes can result in the loss of crucial low- or high-level information from the original feature maps. In this paper, we introduce the Octave C2f module, which employs octave convolution to seamlessly fuse feature maps of different sizes without the need for additional processing. This innovative approach enhances accuracy while reducing computational complexity. Experimental results on the PASCAL VOC and MS COCO datasets demonstrate improved accuracy, reduced computational effort, and a decrease in parameter count compared to the default YOLOv8 model.  
摘要:  
최근 객체 탐지 분야에서는 다양한 크기의 객체를 감지하기 위하여 여러 스케일의 특징 맵들을 융합하는 다중 스케일 특징 융합을 위한 많은 연구가 진행되고 있다. 이러한 다중 스케일의 특징맵을 융합하기 위해, FPN(Feature Pyramid Network)이나 PANet(Path Aggregation Network) 등의 다양한 네트워크 구조가 제시되었다. FPN을 개선한 PANet은 하향식 경로뿐만 아니라 상향식 경로를 추가하여 객체 탐지 분야에서 큰 성능 향상을 이루었다. 그러나 기존 PANet에서의 다중 스케일 특징 융합을 위한 업스케일링 또는 다운스케일링 과정은 원래 특징 맵에서 보존된 저수준 또는 고수준 정보의 손실을 야기했다. 본 논문에서는 옥타브 합성곱(Octave Convolution)을 통한 별도의 처리 없이 다양한 크기의 특징 맵을 원활하게 융합할 수 있는 Octave C2f 모듈을 제안하여 정확성을 향상시키고 계산 복잡도를 줄였다. PASCAL VOC 및 MS COCO 데이터 세트를 사용한 실험 결과, YOLOv8 기본 모델과 비교하여 개선된 정확성과 감소된 계산 작업량 및 매개변수 개수를 확인할 수 있었다.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2383-630X  
入藏号:  
KJD:ART003117142  
Indexed Date:  
2024-09-25

### 41. Semisupervised Semantic Segmentation with Mutual Correction Learning

摘要: The semisupervised semantic segmentation method uses unlabeled data to effectively reduce the required labeled data, and the pseudo supervision performance is greatly influenced by pseudo labels. Therefore, we propose a semisupervised semantic segmentation method based on mutual correction learning, which effectively corrects the wrong convergence direction of pseudo supervision. The well-calibrated segmentation confidence maps are generated through the multiscale feature fusion attention mechanism module. More importantly, using internal knowledge, a mutual correction mechanism based on consistency regularization is proposed to correct the convergence direction of pseudo labels during cross pseudo supervision. The multiscale feature fusion attention mechanism module and mutual correction learning improve the accuracy of the entire learning process. Experiments show that the MIoU (mean intersection over union) reaches 75.32%, 77.80%, 78.95%, and 79.16% using 1/16, 1/8, 1/4, and 1/2 labeled data on PASCAL VOC 2012. The results show that the new approach achieves an advanced level.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1687-5265  
eISSN:  
1687-5273  
入藏号:  
WOS:000884397500010  
PubMed ID:  
36225546  
Indexed Date:  
2022-11-25  
Institution Address:  
Dalian Univ, Key Lab Adv Design & Intelligent Comp, Minist Educ, Dalian 116622, Peoples R China  
Institution Address:  
Dalian Univ Technol, Sch Comp Sci & Technol, Dalian 116024, Peoples R China  
Institution:  
Institution:

### 42. Dual-Attention-Guided Multiscale Feature Aggregation Network for Remote

摘要: Remote sensing image change detection plays an important role in urban planning and environmental monitoring. However, the existing change detection algorithms have limited ability in feature extraction, feature relationship understanding, and capture of small target features and edge detail features, which leads to the loss of some edge detail information and small target features. To this end, a new dual-attention-guided multiscale feature aggregation network is proposed. In the encoding stage, the fully convolutional dual-branch structure is used to extract the semantic features of different scales, and then, the multiscale adjacent semantic information aggregation module is used to aggregate the adjacent semantic features at different scales, which can better capture and fuse the features of different scales, thereby improving the accuracy and robustness of change detection. In the decoding stage, the dual-attention fusion module is proposed to guide and fuse the features extracted from different scales along the spatial and channel directions and reduce the background noise interference. In addition, this article also proposes a three-branch feature fusion module and a global semantic information enhancement module to make the network better integrate global semantics and differential semantics and further integrate high-level semantic features. We also introduce an auxiliary classifier in the decoding stage to provide additional supervision signals and fuse the output of the three auxiliary classifiers with the output of the main decoder to further achieve multiscale feature fusion. The comparative experiments on three remote sensing datasets show that the proposed method is superior to the existing change detection methods.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
28  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
28  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001262073400001  
Indexed Date:  
2024-07-14  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Collaborat Innovat Ctr Atmospher Environm & Equip, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Forestry Univ, Coll Informat Sci & Technol, Nanjing 210037, Peoples R China  
Institution:  
Institution:

### 43. Lightweight Feature Fusion Network for Object Detection in Aerial

摘要: The existing aerial photography image object detection algorithms have several problems, such as complicated models, too many hyperparameters, and poor detection accuracy. Therefore, this paper proposes a lightweight multiscale feature fusion network for object detection in aerial photography images. The proposed network employs the idea of Anchor-Free and reduces the hyperparameters related to Anchor through pixel-by-pixel prediction. First, MobileNetV3 is adopted as the backbone network for feature extraction, and the Ghost bottleneck module is used as the base block for multiscale feature fusion to reduce number of parameters and computational costs. Then, deformable convolution is introduced to construct a deformable receptive field block to improve the robustness of the detector to the deformation of aerial photography objects. Furthermore, the label assignment strategy SimOTA is employed for dynamic sample matching, which alleviates the problems of dense distribution and heavy occlusion of aerial photography objects. The proposed network is evaluated on VisDrone2019-DET and NWPU VHR-10 datasets. The detection accuracy AP50 of the proposed network reaches 26. 6% and 94. 4%, and the detection speed reaches 59. 9 and 79. 6 frame/s, respectively. Compared with other mainstream object detection networks, the proposed network has fewer parameters and computational costs while maintaining high detection accuracy and speed, making it more suitable for airborne computing devices.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1006-4125  
入藏号:  
WOS:000996117100029  
Indexed Date:  
2023-06-14  
Institution Address:  
Tianjin Univ, Sch Microelect, Tianjin 300072, Peoples R China  
Institution Address:  
Phytium Technol Co Ltd, Tianjin 300459, Peoples R China  
Institution Address:  
Tianjin Key Lab Imaging & Sensing Microelect Tech, Tianjin 300072, Peoples R China  
Institution:  
Institution:  
Phytium Technol Co Ltd  
Institution:  
Tianjin Key Lab Imaging & Sensing Microelect Tech

### 44. Road target detection algorithm for autonomous driving scenarios based

摘要: When vehicles, pedestrians, bicycles, and other targets are detected in complex road scenes, the existence of multiscale targets and partial occlusions may easily cause missed and false detections. In this paper, a road target detection algorithm is proposed based on improved YOLOv5s, orienting to autonomous driving scenarios. First, depthwise separable convolution is used to replace partial ordinary convolutions to reduce the number of parameters of the model to improve the detection speed. An improved RFB-s based on receptive field block (RFB) is introduced into the feature fusion network to enhance the effective receptive field area of the feature map, improving the network feature expression capability and the recognizability of the target features by imitating human visual perception. Finally, an adaptive spatial feature fusion method is used to enhance the effect of PANet on multiscale feature fusion. The experimental results reveal that, on the PASCAL VOC dataset, compared with YOLOv5s, the mean value of the average detection precision of the proposed algorithm is improved by 1.71%, reaching 84.01%. Under the premise of meeting the real-time requirement of autonomous driving vehicles, this algorithm has reduced false and missed detections in the target detection to a certain extent, effectively improving the detection performance of the model in complex driving scenarios.  
摘要:  
在复杂道路场景中检测车辆、行人、自行车等目标时,存在因多尺度目标及部分遮挡易造成漏检及误检等情况,提出一种基于改进YOLOv5s的面向自动驾驶场景的道路目标检测算法。首先,利用深度可分离卷积替换部分普通卷积,减少模型的参数量以提升检测速度。其次,在特征融合网络中引入基于感受野模块(receptive field block,RFB)改进的RFB-s,通过模仿人类视觉感知,增强特征图的有效感受野区域,提高网络特征表达能力及对目标特征的可辨识性。最后,使用自适应空间特征融合(adaptively spatial feature fusion,ASFF)方式以提升PANet对多尺度特征融合的效果。实验结果表明,在PASCAL VOC数据集上,所提算法检测平均精度均值相较于YOLOv5s提高1.71个百分点,达到84.01%,在满足自动驾驶汽车实时性要求的前提下,在一定程度上减少目标检测时的误检及漏检情况,有效提升模型在复杂驾驶场景下的检测性能。  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1673-4785  
入藏号:  
CSCD:7802010  
Indexed Date:  
2024-11-15  
Institution Address:  
中国民航大学机器人研究所, 天津 300300, 中国  
Institution Address:  
Robotics Institute,Civil Aviation University of China, Tianjin 300300, China  
Institution:  
中国民航大学机器人研究所  
Institution:  
Robotics Institute,Civil Aviation University of China

### 45. Semantic Correlation Attention-Based Multiorder Multiscale Feature

摘要: Human motion prediction is to predict future human states based on the observed human states. However, current research ignores the semantic correlations between body parts (joints and bones) in the observed human states and motion time; thus, the prediction accuracy is limited. To address this issue, we propose a novel semantic correlation attention-based multiorder multiscale feature fusion network (SCAFF), which includes an encoder and a decoder. In the encoder, a multiorder difference calculation module (MODC) is designed to calculate the multiorder difference information of joint and bone attributes in the observed human states. Then, multiple semantic correlation attention-based graph calculation operators (SCA-GCOs) are stacked to extract the multiscale features of the multiorder difference information. Each SCA-GCO captures joint and bone dependencies of the multiorder difference information, refines them with a semantic correlation attention module (SCAM), and captures temporal dynamics of the refined joint and bone dependencies as the output features. Note that SCAM learns a semantic attention mask describing the semantic correlations between body parts and motion time for feature refinement. Afterward, multiple multiorder feature fusion modules (MOFFs) and multiscale feature fusion modules (MSFFs) are designed to fuse the multiscale features of the multiorder difference information extracted by multiple SCA-GCOs, thus obtaining the motion features of the observed human states. Based on the obtained motion features, the decoder recurrently recruits a composite gated recurrent module (CGRM) and multilayer perceptrons (MLPs) to predict future human states. As far as we know, this is the first attempt to consider the semantic correlations between body parts and motion time in human motion prediction. The results on public datasets demonstrate that SCAFF outperforms existing models.  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
2168-2267  
eISSN:  
2168-2275  
入藏号:  
WOS:000829071600001  
PubMed ID:  
35839189  
Indexed Date:  
2022-07-29  
Institution Address:  
Cent South Univ, Sch Automat, Changsha 410083, Peoples R China  
Institution Address:  
Changsha Hisense Intelligent Syst Res Inst Co Ltd, Res & Dev Dept, Changsha 410208, Peoples R China  
Institution:  
Institution:  
Changsha Hisense Intelligent Syst Res Inst Co Ltd

### 46. No-reference image quality assessment based on multiscale feature

摘要: The no-reference image quality assessment (NR-IQA) method can evaluate the distortions in an image without the reference image. However, due to the diversity of the image contents and distortion types, it is hard for the existing NR-IQA algorithms to obtain competitive performance on both synthetically and authentically distorted images. To address the problem, a multiscale feature representation-based NR-IQA method that performs well for both synthetic and authentic distortions is proposed. This model consists of two parts: The feature extraction part and the feature fusion part. First, part of the Res2Net-50 network is chosen as the feature extraction part due to its high ability in increasing the range of receptive fields. Then, the feature fusion part consisting of a novel residual block and two fully connected layers is designed to fuse the extracted features and realize the quality score mapping. After a series of stepwise optimization experiments, the most competitive network architecture consisting of the feature extraction part and the feature fusion part is obtained. Comprehensive experiments on the LIVE, TID2013, CSIQ, KADID-10k, KonIQ-10K, and LIVE challenge databases demonstrate that the proposed method can work powerfully on both the synthetic and authentic distortions and also has a strong generalization ability.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1751-9659  
eISSN:  
1751-9667  
入藏号:  
WOS:000691410200001  
Indexed Date:  
2021-09-07  
Institution Address:  
Northeast Normal Univ, Sch Phys, Changchun 130024, Peoples R China  
Institution:

### 47. Semantic Segmentation of Remote-Sensing Images Based on Multiscale

摘要: In recent years, the automatic extraction of remote-sensing image information has attracted full attention. However, the particularity of remote-sensing images and the scarcity of data sets with label information have brought new challenges to existing methods. Therefore, we develop a lightweight semantic segmentation network based onmultiscale feature fusion (MFF) and attention refinement (MFFANet). Our network relies on three crucial modules for improved performance. The multiscale attention refinement module strengthens the representation ability of feature maps extracted by the deep residual network. The MFF module aggregates the information carried by the high-level and low-level features while restoring the image resolution. Furthermore, the boundary enhancement module captures boundary details to solve the semantic ambiguity problem. We achieve 83.5% mean intersection over union (MIoU) on the Urban Semantic 3-D (US3D) data set and 69.3% MIoU on the Vaihingen data set with only 8.2M parameters.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
被引频次合计:  
11  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000732210700001  
Indexed Date:  
2022-01-01  
Institution Address:  
China Univ Min & Technol, Sch Comp Sci & Technol, Xuzhou 221116, Jiangsu, Peoples R China  
Institution Address:  
Minist Educ Peoples Republ China, Engn Res Ctr Mine Digitizat, Xuzhou 221116, Jiangsu, Peoples R China  
Institution Address:  
Disaster Intelligent Prevent & Control & Emergenc, Xuzhou 221116, Jiangsu, Peoples R China  
Institution Address:  
QianXuesen Lab Space Technol, Beijing 100094, Peoples R China  
Institution:  
Institution:  
Minist Educ Peoples Republ China  
Institution:  
Disaster Intelligent Prevent & Control & Emergenc  
Institution:  
QianXuesen Lab Space Technol

### 48. Lightweight Infrared Small Target Detection Network Using Full-Scale

摘要: In order to improve detection performance in a U-net-based IR small target detection (IRSTD) algorithm, it is crucial to fuse low- and high-level features. Conventional algorithms perform feature fusion by adding a convolution layer to the skip pathway of the U-net and by connecting the skip connection densely. However, with the added convolution operation, the number of parameters of the network increases, hence the inference time increases accordingly. Therefore, in this letter, a UNet3+-based full-scale skip connection U-net is used as a base network to lower the computational cost by fusing the feature with a small number of parameters. Moreover, we propose an effective encoder and decoder structure for improved IRSTD performance. A residual attention block is applied to each layer of the encoder for effective feature extraction. As for the decoder, a residual attention block is applied to the feature fusion Section to effectively fuse the hierarchical information obtained from each layer. In addition, learning is performed through full-scale deep supervision to reflect all the information obtained from each layer. The proposed algorithm, coined attention multiscale feature fusion U-net (AMFU-net), can hence guarantee effective target detection performance and a lightweight structure [mean intersection over union (mIoU)]: 0.7512 and frame per seconds (FPS): 86.1). Pytorch implementation is available at: github.com/cwon789/AMFU-net.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
20  
被引频次合计:  
20  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001000622900004  
Indexed Date:  
2023-06-17  
Institution Address:  
Seoul Natl Univ, Dept Aerosp Engn & Automat, Seoul 08826, South Korea  
Institution Address:  
Seoul Natl Univ, Syst Res Inst, Seoul 08826, South Korea  
Institution Address:  
Seoul Natl Univ, Dept Aerosp Engn, Nav & Elect Syst Lab, Seoul 08826, South Korea  
Institution Address:  
Seoul Natl Univ, Inst Adv Aerosp Technol, Seoul 08826, South Korea  
Institution:  
Institution:  
Institution:  
Institution:

### 49. Fault Diagnosis Method of the USV Based on Improved Res2Net Multiscale

摘要: With the application of unmanned surface vessel (USV) in military and civilian fields, people have put forward higher requirements for the safety and reliability of USV. When the USV fails, its autonomy will be affected and may cause the USV unable to return back. It is important to accurately detect the faults of USV to ensure their safe operation. In response to the problem of difficult extraction of fault features for USV in complex water conditions, an improved Res2Net multiscale feature fusion network based on the Gramian angular field (GAF) is proposed. First, the vibration signal is converted to GAF. Second, it is input into the constructed depthwise separable inverted Res2Net residual block (DSIRes2Net) to extract the multiscale features. The multiscale features are integrated into the low-dimensional features. Finally, these features are put into the adaptive time-frequency feature fusion block (ATFF) to adaptively fuse the time-domain and frequency-domain features, so as to realize the fault diagnosis of the USV. The experimental results show that the fault diagnosis accuracy of the USV obtained from the proposed algorithm can reach 98.5%, with a variance of 0.17 and a standard deviation of 0.41. The accuracy and stability of the fault diagnosis model have been effectively improved. It provides an important guarantee for the safe operation of the USV.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001285340000134  
Indexed Date:  
2024-09-27  
Institution Address:  
Shandong Univ Sci & Technol, Coll Elect Engn & Automat, Qingdao 266590, Peoples R China  
Institution Address:  
Shandong Univ Sci & Technol, Coll Safety Sci & Engn, Qingdao 266590, Peoples R China  
Institution:  
Institution:

### 50. Image Segmentation of Retinal Blood Vessels Based on Dual-Attention

摘要: Aiming at the problem of insufficient details of retinal blood vessel segmentation in current research methods, this paper proposes a multiscale feature fusion residual network based on dual attention. Specifically, a feature fusion residual module with adaptive calibration weight features is designed, which avoids gradient dispersion and network degradation while effectively extracting image details. The SA module and ECA module are used many times in the backbone feature extraction network to adaptively select the focus position to generate more discriminative feature representations; at the same time, the information of different levels of the network is fused, and long-range and short-range features are used. This method aggregates low-level and high-level feature information, which effectively improves the segmentation performance. The experimental results show that the method in this paper achieves the classification accuracy of 0.9795 and 0.9785 on the STARE and DRIVE datasets, respectively, and the classification performance is better than the current mainstream methods.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
1748-670X  
eISSN:  
1748-6718  
入藏号:  
WOS:000829462800006  
PubMed ID:  
35844462  
Indexed Date:  
2022-07-30  
Institution Address:  
Henan Univ Engn, Sch Comp, Zhengzhou 451191, Peoples R China  
Institution Address:  
Henan Univ Engn, Sch Elect Informat Engn, Zhengzhou 451191, Peoples R China  
Institution:  
Institution:

### 51. PFDN: A Polarimetric Feature-Guided Deep Network for Dual-Polarized SAR

摘要: As one important application of synthetic aperture radar (SAR), ship classification attracts researchers' attention in recent years. To improve the accuracy of ship classification in dual-polarized SAR images, we here put forward a novel polarimetric feature-guided deep network (PFDN). Detailedly, a new polarimetric feature smoothed-polarimetric information fusion (SPF) is first built through fusing both amplitudes of dual-polarized channels. Since only the amplitude information are used, SPF cannot be affected by the phase noise. Then, another two key components, i.e., the multiscale feature fusion attention module (MFFAM) and the dynamic gating feature fusion mechanism (DGFFM), are further proposed to extract deeper classification features. Finally, via combining these three different modules together, PFDN is constructed. Experiments tested on the dataset OpenSARShip2.0 show that PFDN can achieve higher accuracies (87.13% in three-class task and 65.97% in six-class task) than the other state-of-the-art (SOTA) methods.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001178137500005  
Indexed Date:  
2024-04-26  
Institution Address:  
Shanghai Jiao Tong Univ, Sch Sensing Sci & Engn, Shanghai Key Lab Intelligent Sensing & Recognit, Shanghai 200240, Peoples R China  
Institution:

### 52. Face Recognition Algorithm Based on Multiscale Feature Fusion Network

摘要: A face recognition model based on a multiscale feature fusion network is constructed, aiming to make full use of the characteristics of face and to improve the accuracy of face recognition. In addition, three different scale networks are designed to extract global features of faces. Multiscale cross-layer bilinear features of multiple networks are integrated via introducing a hierarchical bilinear pooling layer. By capturing some of the feature relationships between different levels, the model's ability to extract and distinguish subtle facial features is enhanced. Simultaneously, this study uses layer-by-layer deconvolution to fuse multilayer feature information, to solve the problem of losing some key features when extracting features from multilayer convolutional layers and pooled layers. The experimental results show that compared with the recognition accuracy of traditional algorithms, the recognition accuracy of the algorithm on Yale, AR, and ORL face databases is significantly improved.  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
1687-5265  
eISSN:  
1687-5273  
入藏号:  
WOS:000783742700007  
PubMed ID:  
35341167  
Indexed Date:  
2022-04-29  
Institution Address:  
Jiaozuo Normal Coll, Sch Informat Engn, Jiaozuo 454000, Henan, Peoples R China  
Institution:  
Jiaozuo Normal Coll

### 53. A Defect Recognition Method for Low-Quality Weld Image Based on

摘要: Defect recognition is an essential technology for the safety of pipelines. Recently, deep-learning (DL)-based methods have shown great progress in recognizing pipeline weld defects. However, when it comes to low-quality images in some complicated applications, the DL-based methods can hardly achieve satisfactory results. To solve this problem, this article proposes a consistent multiscale feature mapping method for defect recognition of low-quality weld images. First, a multiscale feature mapping method is proposed, so that different types of defect images can be mapped into distinctive features in both local and global embedding spaces. Second, a novel consistency strategy is proposed to maintain the consistency of the local and global embedding spaces, so that similar defects can be easily distinguished. Third, a special feature fusion model is proposed to fuse the mapped local and global features, so that the performance of defect recognition can be improved. Finally, five groups of experiments are conducted using pipeline weld X-ray images and metal weld X-ray images. The experiment results show that the proposed method is effective in recognizing the weld defects of low-quality X-ray images.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
13  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
13  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000799630900004  
Indexed Date:  
2022-06-03  
Institution Address:  
Northeastern Univ, State Key Lab Synthet Automat Proc Ind, Shenyang 110819, Peoples R China  
Institution Address:  
Northeastern Univ, Coll Informat Sci & Engn, Shenyang 110819, Peoples R China  
Institution Address:  
Univ Sci & Technol Beijing, Sch Civil & Resource Engn, Beijing 100083, Peoples R China  
Institution Address:  
Shenyang Paidelin Technol Co Ltd, Shenyang 110081, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Shenyang Paidelin Technol Co Ltd

### 54. Traffic Sign Detection Algorithm Based on Modified Anchor-Free Model

摘要: Traffic sign detection is an essential function of autonomous driving systems, and most modern traffic sign detectors are anchor-based, traversing potential object locations based on anchors. To solve the problems of heavy computing costs and the need to set several hyperparameters in anchor-based models, we propose an anchor-free traffic sign detection algorithm based on an encoder-decoder structure. We introduce a residual augmentation branch in the decoder module in this study to improve feature expression ability during the decoding process. To improve the ability to detect multiscale traffic signs, we propose a multiscale feature fusion subnetwork to effectively extract and use multiscale features. A Ghost lightweight module is adopted by the multiscale feature extraction module, which indistinctively increases the computational cost. On the Tsinghua-Tencent 100 K dataset, our approach achieved a recall of 92.5% and an accuracy of 90.3%, while the model's parameter amount and model size are approximately 1.61\*10~7 and 64.4 Mbit, respectively. The experimental results show that the proposed algorithm outperforms the mainstream object detection algorithms in terms of precision, computing cost, and overall performance.  
摘要:  
交通标志检测是自动驾驶系统的一项重要功能,当前先进的交通标志检测器大多采用Anchor-Based网络模型,根据锚框遍历所有潜在的目标位置。为了减少锚框带来的计算开销和过多的超参数设置,提出了一种基于编码-解码结构的Anchor-Free交通标志检测算法。为了增加解码模块的特征表征能力,在解码模块中引入残差增强分支。为了高效地提取和利用多尺度特征,设计了特征融合子网络,提升对多尺度目标的检测能力,并使用Ghost轻量化模块提取多尺度特征图,不显著引进运算量。在Tsinghua-Tencent 100K数据集上进行验证,所提算法实现了92.5%的召回率和90.3%的准确率,模型的参数量和模型大小分别为1.61\*10~7和64.4 Mbit。实验结果表明,与主流目标检测算法相比,所提算法的检测精度较高,计算开销较低,在综合性能上具有优越性。  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7384667  
Indexed Date:  
2023-05-05  
Institution Address:  
天津大学电气自动化与信息工程学院, 天津 300072, 中国  
Institution Address:  
School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China  
Institution:  
天津大学电气自动化与信息工程学院  
Institution:  
School of Electrical and Information Engineering, Tianjin University

### 55. A Method for Surface Defect Detection Based on Multiscale Feature Fusion

摘要: The two-stage defect detection model needs to pay attention to the results of the segmentation network and the classification network, and the results of the segmentation network will have an impact on the classification network. Previous models ignored shallow features in the segmentation network and used relatively simple classification networks that could not make good use of the features of the segmentation network. This paper proposes a surface defect detection algorithm based on multi-scale feature fusion and pyramid attention(MFFPA). First, a multi-scale feature fusion module is added to the segmentation network to fuse shallow features and extract more comprehensive feature information; then a pyramid attention module is added to the classification network to increase the receptive field of the model and enhance the discriminative ability of the model. The method proposed in this article was verified on four datasets, and the experimental results show that the added module can effectively improve the accuracy of the model.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001184738000001  
Indexed Date:  
2024-03-31  
Institution Address:  
Changzhou Univ, Sch Comp Sci & Artificial Intelligence, Changzhou 213000, Peoples R China  
Institution:

### 56. Global-Local Collaborative Enhancement-Based Remote Sensing Target

摘要: To enhance detection performance for various targets in remote sensing images, an improved algorithm based on LSKNet-S is proposed. This algorithm notably advances the network's ability to detect remote sensing targets by integrating features across diverse receptive field levels. First, a multiscale feature fusion module is designed to strengthen the model's ability to extract global contextual information, with improvements to the multilayer perceptron. Concurrently, a lightweight local visual center module is introduced, enhancing the model's sensitivity to local features. The integration of these modules facilitates effective multiscale feature extraction and fusion within the model. Additionally, a scale-enhancing upsampling operation is incorporated within the detection head, which elevates the feature map resolution, allowing the model to more effectively capture detailed information on various targets within remote sensing images. Experimental results indicate that the proposed algorithm improves the mean average precision (mAP) by 3. 43 percentage points on the HRSC2016 dataset and by 1. 12 percentage points on the DIOR-R dataset, outperforming current mainstream algorithms. These results confirm the effectiveness of the proposed algorithm in remote sensing object detection contexts.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001438407600013  
Indexed Date:  
2025-03-13  
Institution Address:  
Tianjin Chengjian Univ, Coll Comp & Informat Engn, Tianjin 300384, Peoples R China  
Institution:

### 57. Remote Sensing Image-Matching Network Based on Multiscale Feature Fusion

摘要: Remote sensing image matching is one of the fundamental challenges in earth observation. The complexity and diversity of surface information in remote sensing images often pose difficulties for image matching. To overcome these difficulties, a remote sensing image-matching network based on multiscale feature fusion and importance ranking loss is proposed. This network comprises two parts: a key-point detection network and a feature descriptor extraction network. The key-point detection network has a multilayer convolutional structure based on feature pyramids. This structure is designed to achieve multiscale feature fusion at different network levels. Multiple convolution kernels are used to gradually expand receptive fields at the same level, thereby fully capturing multiscale information in remote sensing images. Furthermore, CBAM is used to aggregate the response graph of the key-point detection network to detect key points with significant scores. The key-point detection network is optimized using the score loss and image block loss, and the feature descriptor sub-extraction network is optimized using the descriptor subloss. The score-importance sorting loss function, descriptor sub-importance sorting loss function, and neighbor mask-based descriptor subloss function are specially designed to ensure that the key points, descriptors, and image blocks used for remote sensing image matching have high repeatability and distinguishability, which improves the accuracy of remote sensing image matching. In this study, many remote sensing images were collected and a remote sensing image-matching dataset was constructed via homography transformation. This dataset was used to experimentally verify the performance of the proposed network model. Compared with traditional image-matching methods or other end-to-end deeplearning image-matching methods, the proposed network model has considerable advantages in remote sensing image matching.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001395792100012  
Indexed Date:  
2025-02-12  
Institution Address:  
Hebei Univ Technol, Sch Artificial Intelligence, Tianjin 300401, Peoples R China  
Institution:

### 58. An End-to-End Defect Detection Method for Mobile Phone Light Guide Plate

摘要: Automatic vision-based defect detection on the mobile light guide plate (LCP) is a challenging task due to the low contrast between the defect and the background, uneven brightness, and complex gradient texture. An end-to-end multitask learning network architecture for the defect detection of mobile phone I.GP is proposed. First, the main structure of the multitask learning network is designed. The encoder part uses a similar U-Net encoder structure to obtain multiscale features, and the feature fusion part adopts feature fusion to interact with multiscale features. Second, the segmentation head is designed to complete the precise location of each defect in an image by using the multiscale feature fusion, which prepares it for the quantification of defect characteristics. Combining the multiscale features and the output mask of the segmentation head, the classification head is designed to accurately detect the defects of mobile phone LGP. Finally, the defect detection data set has been constructed based on the mobile phone LGP images collected on the industrial site, and a lot of experiments are performed on the mobile phone LGP data set and Kolektor surface-defect data set (KolektorSDD). The experimental results show that the Fl-score on the two data sets can reach 99.67% and 96.77%, respectively, which verifies the effectiveness of the method proposed in this article.  
Web of Science 核心合集中的 "被引频次":  
30  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
30  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000616309500017  
Indexed Date:  
2021-03-05  
Institution Address:  
Zhejiang Sci Tech Univ, Sch Fac Mech Engn & Automat, Hangzhou 310000, Peoples R China  
Institution:

### 59. Lightweight Feature Fusion Network for Object Detection in Aerial

摘要: The existing aerial photography image object detection algorithms have several problems, such as complicated models, too many hyperparameters, and poor detection accuracy. Therefore, this paper proposes a lightweight multiscale feature fusion network for object detection in aerial photography images. The proposed network employs the idea of Anchor-Free and reduces the hyperparameters related to Anchor through pixel-by-pixel prediction. First, MobileNetV3 is adopted as the backbone network for feature extraction, and the Ghost bottleneck module is used as the base block for multiscale feature fusion to reduce number of parameters and computational costs. Then, deformable convolution is introduced to construct a deformable receptive field block to improve the robustness of the detector to the deformation of aerial photography objects. Furthermore, the label assignment strategy SimOTA is employed for dynamic sample matching, which alleviates the problems of dense distribution and heavy occlusion of aerial photography objects. The proposed network is evaluated on VisDrone2019-DET and NWPU VHR-10 datasets. The detection accuracy AP50 of the proposed network reaches 26.6% and 94.4%, and the detection speed reaches 59.9 and 79.6 frame/s, respectively. Compared with other mainstream object detection networks, the proposed network has fewer parameters and computational costs while maintaining high detection accuracy and speed, making it more suitable for airborne computing devices.  
摘要:  
针对现有航摄图像目标检测算法中模型复杂、超参数多、检测精度较低的问题,提出一种面向航摄图像目标检测的轻量级多尺度特征融合网络。该网络采用Anchor-Free思想,通过逐像素预测的方式,减少了与Anchor相关的超参数;利用MobileNetV3作为特征提取网络并使用Ghost瓶颈模块优化多尺度特征融合网络,来降低网络的参数量和计算量;引入可变形卷积来构建可变形感受野模块,提高检测器对航摄图像目标形变的鲁棒性;同时采用标签分配策略SimOTA进行动态样本匹配,以缓解航摄图像目标分布密集、遮挡严重的检测问题。在数据集VisDrone2019-DET和NWPU VHR-10上对所提网络进行评估,检测精度AP50分别达26.6%和94.4%,检测速度分别达59.9 frame/s和79.6 frame/s。与主流目标检测网络相比,所提网络在保持较高检测精度和速度的同时,具有较小的参数量和计算量,更适合应用于机载计算设备。  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7495855  
Indexed Date:  
2023-09-28  
Institution Address:  
天津大学微电子学院, 天津 300072, 中国  
Institution Address:  
天津大学微电子学院;;天津市成像与感知微电子技术重点实验室, ;;天津市成像与感知微电子技术重点实验室, ;;, 天津;;天津 300072;;300072, 中国  
Institution Address:  
飞腾信息技术有限公司, 天津 300459, 中国  
Institution Address:  
School of Microelectronics, Tianjin University, Tianjin 300072, China  
Institution Address:  
School of Microelectronics, Tianjin University;;Tianjin Key Laboratory of Imaging and Sensing Microelectronic Technology, ;;Tianjin Key Laboratory of Imaging and Sensing Microelectronic Technology, ;;, Tianjin;;Tianjin 300072;;300072  
Institution Address:  
Phytium Technology Co., Ltd., Tianjin 300459, China  
Institution:  
天津大学微电子学院  
Institution:  
天津大学微电子学院;;天津市成像与感知微电子技术重点实验室  
Institution:  
飞腾信息技术有限公司  
Institution:  
School of Microelectronics, Tianjin University  
Institution:  
School of Microelectronics, Tianjin University;;Tianjin Key Laboratory of Imaging and Sensing Microelectronic Technology  
Institution:  
Phytium Technology Co., Ltd.

### 60. Improved PointPillar point cloud object detection based on feature

摘要: A point cloud object detection network, Pillar-FFNet, is proposed by introducing a multiscale feature fusion strategy and an attention mechanism to address the ineffectiveness of PointPillar in detecting small sparse objects in point clouds in autonomous driving road scenarios. First, a backbone network based on a residual structure is designed for feature extraction in the network. Second, a simple and effective multiscale feature fusion strategy is designed to address the problem that the feature maps fed into the detection head do not make full use of the semantic information of high-level features and the spatial information of low-level features. Finally, a convolutional attention mechanism is proposed to treat information redundancy in the feature maps extracted using the backbone network. To validate the performance of the proposed algorithm, experiments are conducted on the KITTI and DAIR-V2X-I datasets. The results show that the proposed algorithm achieves maximum average accuracy improvements of 0.84%, 2.13%, and 4.02% for cars, pedestrians, and cyclists, respectively, on the KITTI dataset and maximum average accuracy improvements of 0.33%, 2.09%, and 4.71% for cars, pedestrians, and cyclists, respectively, on the DAIR-V2X-I dataset compared with the PointPillar results. Experimental results demonstrate the effectiveness of the proposed method for the detection of sparse small objects in point clouds.  
摘要:  
针对PointPillar在自动驾驶道路场景下对点云稀疏小目标检测效果差的问题,通过引入一种多尺度特征融合策略和注意力机制,提出一种点云目标检测网络Pillar-FFNet。针对网络中的特征提取问题,设计了一种基于残差结构的主干网络;针对馈入检测头的特征图没有充分利用高层特征的语义信息和低层特征的空间信息的问题,设计了一种简单有效的多尺度特征融合策略;针对主干网络提取的特征图中信息冗余的问题,提出了一种卷积注意力机制。为验证所提算法的性能,在KITTI和DAIR-V2X-I数据集上进行实验。实验结果表明,所提出的算法在KITTI数据集上与PointPillar相比,汽车、行人和骑行者的平均精度最大提高分别为0.84%,2.13%和4.02%;在DAIR-V2X-I数据集上与PointPillar相比,汽车、行人和骑行者的平均精度最大提高分别为0.33%,2.09%和4.71%,由此证明了所提方法对点云稀疏小目标检测的有效性。  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
2  
ISSN:  
1004-924X  
入藏号:  
CSCD:7568538  
Indexed Date:  
2023-12-29  
Institution Address:  
国防科技大学电子科学学院, ATR国家重点实验室, 长沙, 湖南 410073, 中国  
Institution Address:  
College of Electronic Science and Technology, National University of Defense Technology, National Key Laboratory of Science and Technology on Automatic Target Recognition, Changsha, Hunan 410073, China  
Institution:  
国防科技大学电子科学学院  
Institution:  
College of Electronic Science and Technology, National University of Defense Technology

### 61. UAV detection in complex background with multi-scale feature fusion

摘要: The reckless flight of unmanned aerial vehicle (UAV) seriously threatens the public and aviation safety. Due to their small size and unobvious features, it remains a great challenge for the current detection algorithms to detect UAV, especially in complex backgrounds with backlighting. To address these issues, the multiscale feature fusion enhancement strategy and channel-weight matching (CWM) rule are proposed in this paper. A multiscale feature fusion enhancement strategy is presented to capture the multi-scale contextual information, which not only suppresses information conflicts but also enhances feature extraction capabilities. Then, an up-sampling method based on CWM is designed to enhance the sensitivity of small object, which uses different up-sampling techniques based on the importance level of each feature channel. Finally, a feature refinement module for small object is designed to further enhance the characterization of their features. The ablation and comparative experiments are carried out on the self-made UAV dataset. Compared to the original YOLOv5 algorithm, the proposed method shows an increase of 3.6% in mAP0.5 and 2.8% in mAP0.5:0.95, respectively. Moreover, the comparative experiments are implemented on the VisDrone2019 dataset, and the results indicate that the mAP0.5 and mAP0.5:0.95 of the proposed method also increase by 4.2% and 1.6%, respectively.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0031-8949  
eISSN:  
1402-4896  
入藏号:  
WOS:001379293700001  
Indexed Date:  
2024-12-22  
Institution Address:  
Beijing Inst Technol, Sch Automat, Beijing, Peoples R China  
Institution Address:  
Yubei Steering Syst Xinxiang Co Ltd, ECU Res Inst, Xinxiang, Peoples R China  
Institution Address:  
Henan Univ Technol, Coll Elect Engn, Zhengzhou, Peoples R China  
Institution:  
Institution:  
Yubei Steering Syst Xinxiang Co Ltd  
Institution:

### 62. MDCC-Net: Multiscale double-channel convolution U-Net framework for

摘要: Purpose: Multiscale feature fusion is a feasible method to improve tumor segmentation accuracy. However, current multiscale networks have two common problems: 1. Some networks only allow feature fusion between encoders and decoders of the same scale. It is obvious that such feature fusion is not sufficient. 2. Some networks have too many dense skip connections and too much nesting between the coding layer and the decoding layer, which causes some features to be lost and means that not enough information will be learned from multiple scales. To overcome these two problems, we propose a multiscale double-channel convolution U-Net (MDCC-Net) framework for colorectal tumor segmentation.  
Methods: In the coding layer, we designed a dual-channel separation and convolution module and then added residual connections to perform multiscale feature fusion on the input image and the feature map after dual channel separation and convolution. By fusing features at different scales in the same coding layer, the network can fully extract the detailed information of the original image and learn more tumor boundary information.  
Results: The segmentation results show that our proposed method has a high accuracy, with a Dice similarity coefficient (DSC) of 83.57%, which is an improvement of 9.59%, 6.42%, and 1.57% compared with nnU-Net, U Net, and U-Net++, respectively.  
Conclusion: The experimental results show that our proposed method has good performance in the segmentation of colorectal tumors and is close to the expert level. The proposed method has potential clinical applicability.  
在中国科学引文数据库中的被引频次:  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
25  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
6  
被引频次合计:  
28  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:000623917000008  
PubMed ID:  
33360107  
Indexed Date:  
2021-03-01  
Institution Address:  
Northeastern Univ, Coll Med & Biol Informat Engn, Shenyang, Peoples R China  
Institution Address:  
Chinese Acad Sci, Shenzhen Inst Adv Technol, Shenzhen, Peoples R China  
Institution Address:  
Harbin Med Univ, Affiliated Hosp 2, Dept Radiol, Harbin, Peoples R China  
Institution Address:  
Hejian Peoples Hosp, Dept Operating Room, Cangzhou, Peoples R China  
Institution Address:  
Guangzhou Med Univ, Dept Surg, Affiliated Hosp 1, Guangzhou, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Hejian Peoples Hosp  
Institution:

### 63. Collaborative Attention Guided Multi-Scale Feature Fusion Network for

摘要: Medical image segmentation is an important and complex task in clinical practices, but the widely used U-Net usually cannot achieve satisfactory performances in some clinical challenging cases. Therefore, some advanced variants of U-Net are proposed using multi-scale and attention mechanisms. Different from the existing works where multi-scale and attention are usually used independently, in this work, we integrate them together and propose a collaborative attention guided multi-scale feature fusion with enhanced convolution based U-Net (EC-CaM-UNet) model for more accurate medical image segmentation, where a novel collaborative attention guided multi-scale feature fusion (CoAG-MuSFu) module is proposed to highlight important (but small and unremarkable) multi-scale features and suppress irrelevant ones in model learning. Specifically, CoAG-MuSF uses a multi-dimensional collaborative attention (CoA) block to estimate the local and global self-attention, which is then deeply fused with the multi-scale feature maps generated by a multi-scale (MuS) block to better highlight the important multi-scale features and suppress the irrelevant ones. Furthermore, an additional supervision path and enhanced convolution blocks are used to enhance the deep model's feature learning ability in both deep and shallow features, respectively. Experimental results on three public medical image datasets show that EC-CaM-UNet greatly outperforms the state-of-the-art medical image segmentation baselines. The codes will be released after acceptance.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
2327-4697  
入藏号:  
WOS:001175221000022  
Indexed Date:  
2024-04-13  
Institution Address:  
Hebei Univ Technol, Sch Hlth Sci & Biomed Engn, State Key Lab Reliabil & Intelligence Elect Equipm, Tianjin 300131, Peoples R China  
Institution Address:  
Hebei Univ Technol, State Key Lab Reliabil & Intelligence Elect Equipm, Tianjin 300131, Peoples R China  
Institution Address:  
Hebei Univ Technol, Sch Artificial Intelligence, Tianjin 300131, Peoples R China  
Institution Address:  
Shenzhen Univ, Coll Comp Sci & Software Engn, Shenzhen 518060, Peoples R China  
Institution Address:  
Shenzhen Univ, Guangdong Lab Artificial Intelligence & Digital Ec, Shenzhen 518060, Peoples R China  
Institution Address:  
Vienna Univ Technol, Inst Log & Computat, A-1040 Vienna, Austria  
Institution Address:  
Univ Oxford, Dept Comp Sci, Oxford OX1 3QD, England  
Institution Address:  
Shenzhen Univ, Coll Comp Sci & Software Engn, Shenzhen 518060, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 64. Discriminative feature fusion for RGB-D salient object detection

摘要: RGB-D salient object detection aims to separate salient object from an image aided by depth. While a number of effective approaches have been proposed, difficulties still exist, which is due to two challenges: (1) It is difficult to fully and effectively fuse RGB and depth features especially in challenging scenes; (2) How to enhance the semantic information of low-level feature and enrich the spatial information of high-level feature. Most of the existing approaches design separate modules to address them. In this paper, a unified discriminative feature fusion module is proposed to be used for both multimodal and multiscale feature fusion. The module can also increase the semantic information in low-level features and enrich the spatial information in high-level features. A multi-scale contextual perception module is embedded in the network to accurately localize objects at different scales. Unlike other methods, the depth branch in the network uses pure convolution for complementary feature extraction. This paper conducted a comparison with 14 state-of-the-art methods on 8 datasets, and the experimental results suggest that the proposed approach is more effective and superior.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0045-7906  
eISSN:  
1879-0755  
入藏号:  
WOS:000960964800001  
Indexed Date:  
2023-04-16  
Institution Address:  
Yangzhou Univ, Sch Informat Engn, Yangzhou, Peoples R China  
Institution Address:  
Yangzhou Hangsheng Technol Co Ltd, Yangzhou, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Fac Sci, Dept Appl Math, Hong Kong, Peoples R China  
Institution:  
Institution:  
Yangzhou Hangsheng Technol Co Ltd  
Institution:

### 65. Attention-Based Multiscale Feature Fusion for Efficient Surface Defect

摘要: Deep-learning-based detection methods have been widely applied to industrial defect inspection. However, directly using vanilla detection methods fails to achieve satisfying performance due to the lack of identifiable features. In this article, a novel attention-based multiscale feature fusion module (AMFF) is proposed, aiming to enhance defect features and improve defect identification by leveraging attention mechanism in the feature fusion. AMFF includes self-enhanced attention module (SEAM) and cross-enhanced attention module (CEAM). SEAM is performed on a single feature map, which first adopts multiple dilation convolutions to enrich contextual information without compromising resolution and then utilizes the intralayer attention on the current feature map. CEAM takes both the current feature map and the adjacent feature map as input to perform cross-layer attention. The adjacent feature map is modulated with the guidance of the current feature map, which is then combined with the current feature map and the output of SEAM for final prediction. AMFF is utilized in current feature fusion networks, e.g., feature pyramid network (FPN) and path aggregation FPN (PAFPN), and is further integrated into prevalent detectors to guide them to pay more attention to defects rather than the background. Extensive experiments are conducted on two real industrial datasets released by Tianchi platform, i.e., fabric and aluminum defect datasets. For each dataset, 500 images are randomly selected for test and the rest for training. The proposed AMFF is demonstrated to significantly boost defect detection accuracy with acceptable computational cost, and the real-time performance could fully satisfy practical requirements.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001188560600008  
Indexed Date:  
2024-04-12  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Multimodal Artificial Intelligence S, Beijing 100190, Peoples R China  
Institution Address:  
Shanghai Tech Univ, Sch Informat Sci & Technol, Shanghai 201210, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Multimodal Artificial Intelligence S, Beijing 100190, Peoples R China  
Institution Address:  
Alibaba Grp Taobao Co Ltd, Hangzhou 311121, Peoples R China  
Institution Address:  
Shanghai Engn Res Ctr Intelligent Vis & Imaging, Shanghai 200050, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Alibaba Grp Taobao Co Ltd  
Institution:  
Shanghai Engn Res Ctr Intelligent Vis & Imaging

### 66. Single shot multibox detector object detection based on attention

摘要: The single shot multibox detector (SSD) is one of the most important algorithms in single-stage target detection, having the characteristics of multiscale detection and rapid detection speed. However, the effective SSD feature layers are independent of one another, which can lead to object detection difficulties. To address this problem, we proposed an improved SSD object detection algorithm. First, the global attention mechanism (GAM)-which can enhance spatial and channel information-was introduced into the multiscale feature layer. The channel attention module of the GAM was improved. Second, a feature fusion module was introduced to strengthen the relationship between feature layers. Finally, the cross stage partial structure was introduced into the feature fusion module, and used to improve the model's learning ability. For model training and detection based on the PASCAL VOC dataset, the mean average precision and frames per second obtained by the improved SSD algorithm were 84.67% and 18.67, respectively, which could effectively detect difficult targets.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:000981645700050  
Indexed Date:  
2023-06-13  
Institution Address:  
Inner Mongolia Univ Technol, Hohhot, Peoples R China  
Institution Address:  
Inner Mongolia Univ Technol, Sch Informat Engn, Hohhot, Peoples R China  
Institution Address:  
Tianjin Univ Technol, Engn, Tianjin, Peoples R China  
Institution:  
Institution:  
Institution:

### 67. AMFF-Net: An Effective 3D Object Detector Based on Attention and

摘要: With the advent of autonomous vehicle applications, the importance of LiDAR point cloud 3D object detection cannot be overstated. Recent studies have demonstrated that methods for aggregating features from voxels can accurately and efficiently detect objects in large, complex 3D detection scenes. Nevertheless, most of these methods do not filter background points well and have inferior detection performance for small objects. To ameliorate this issue, this paper proposes an Attention-based and Multiscale Feature Fusion Network (AMFF-Net), which utilizes a Dual-Attention Voxel Feature Extractor (DA-VFE) and a Multi-scale Feature Fusion (MFF) Module to improve the precision and efficiency of 3D object detection. The DA-VFE considers pointwise and channelwise attention and integrates them into the Voxel Feature Extractor (VFE) to enhance key point cloud information in voxels and refine more-representative voxel features. The MFF Module consists of self-calibrated convolutions, a residual structure, and a coordinate attention mechanism, which acts as a 2D Backbone to expand the receptive domain and capture more contextual information, thus better capturing small object locations, enhancing the feature-extraction capability of the network and reducing the computational overhead. We performed evaluations of the proposed model on the nuScenes dataset with a large number of driving scenarios. The experimental results showed that the AMFF-Net achieved 62.8% in the mAP, which significantly boosted the performance of small object detection compared to the baseline network and significantly reduced the computational overhead, while the inference speed remained essentially the same. AMFF-Net also achieved advanced performance on the KITTI dataset.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001117968500001  
PubMed ID:  
38067692  
Indexed Date:  
2023-12-17  
Institution Address:  
Guangdong Univ Technol, Sch Informat Engn, Guangzhou 510006, Peoples R China  
Institution:

### 68. Natural Scene Text Detection With Multiscale Feature Augmentation and

摘要: Recently, the DB algorithm has drawn considerable attention in scene text detection due to its differentiable binarization module, which is proposed to simplify the complex post-processing of the existing segmentation-based scene text detection approaches. However, DB is limited to its layer-wise multiscale feature representation, semantic information loss for the feature map of the highest level, insufficient localization signals at the higher layers, and less scale robustness of the segmentation model. In this paper, we propose a novel scene text detector with multiscale feature augmentation and attention mechanisms (MFAAM). Specifically, Res2Net serves as the backbone network for extracting fine-grained features with multiple scales. In feature fusion, we construct the Deep Feature Enhancement (DFE) module which can extract ratio-invariant spatial context information to reduce the semantic information loss in the top-level feature map. The Feature Pyramid Augmentation (FPA) module is employed to fuse the lower-level positioning information along the bottom-up path by which the higher-level features are strengthened. Furthermore, the improvement for scale robustness of the segmentation model is implemented by Attentional Feature Fusion (AFF), which can adaptively integrate multiscale features through the attention mechanisms with the channel and spatial context. Experiments on the ICDAR2015 dataset validate the superiority of the proposed method. It is worth noting that the presented detector outperforms DB by 4.38 percent on precision, 2.88 percent on recall, and 3.55 percent on F-measure. Compared with the state-of-the-art algorithms, the presented algorithm exhibits a higher F-measure.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001409550200016  
Indexed Date:  
2025-02-05  
Institution Address:  
Shanxi Univ, Sch Phys & Elect Engn, Taiyuan 030006, Peoples R China  
Institution:

### 69. CL-MFGCN: Graph Structure Contrastive Learning and Multiscale Feature

摘要: To address the conflict between the limited availability of spectrum resources and the swiftly growing number of frequency equipment in the Internet of Vehicles, this article starts from the solution of dynamic access of radio equipment and studies the problem of power spectrum density prediction of different channels, which is called spectrum prediction. We first propose the graph structure learning problem of electromagnetic spectrum data under the graph contrastive learning (GCL) framework from the causal perspective, and establishes a graph structure representation model between channel signal activity rules. Then, we establish the statistical distribution models of different channels at fine resolution based on the Gaussian mixture model. Then, the statistical model is embedded as prior knowledge using a graph convolutional network (GCN), and the channel association features are mapped to the association knowledge embedding using a graph structure encoder. This article proposes a novel spectrum prediction architecture based on GCL and multiscale feature fusion GCN (CL-MFGCN) to mine the time-frequency implicit knowledge of spectrum data, and the above knowledge embedding is integrated. This article visualizes the channel association relationship in the form of a graph structure. The experimental results indicate that the CL-MFGCN, reduces the average MAE by 14.6% (from 1.436 to 1.226) and the average MAPE by 12.9% (from 0.017 to 0.0148) compared to the second-best Pyraformer model, while maintaining a lower model complexity.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2327-4662  
入藏号:  
WOS:001414873100034  
Indexed Date:  
2025-02-14  
Institution Address:  
Harbin Engn Univ, Coll Informat & Commun Engn, Harbin 150000, Peoples R China  
Institution Address:  
Chengdu Fuyuanchen Technol Co Ltd, Chengdu 610000, Peoples R China  
Institution Address:  
Harbin Engn Univ, Coll Comp Sci & Technol, Harbin 150000, Peoples R China  
Institution:  
Institution:  
Chengdu Fuyuanchen Technol Co Ltd  
Institution:

### 70. Semi-supervised single image dehazing based on dual-teacher-student

摘要: While significant progress has been made in image dehazing techniques, the lack of large-scale labeled datasets remains one of the limiting factors for enhancing the performance of image dehazing algorithms. Therefore, based on the mean teacher model, we propose a semi-supervised dehazing framework with a dual-teacher-student (DTS) architecture. DTS is composed of a pretrained teacher network (P-teacher), a mean teacher network (M-teacher), and a student network. The P-teacher facilitates the student network in learning intermediate layer features that resemble haze-free images through knowledge transfer. The M-teacher guides the student network in image dehazing in unsupervised manner. The P-teacher, M-teacher, and student networks share the same network architecture known as the multiscale feature fusion attention-enhanced network (MFFA-Net). The MFFA-Net consists of a multiscale feature fusion network (MFF-Net) and an attention network (A-Net). The MFF-Net is responsible for fusing features from different levels. The A-Net is capable of compensating for information loss during downsampling in the MFF-Net and dynamically adjusting the focus on different regions. Extensive experimental results demonstrate that the dehazing method proposed in this paper outperforms several state-of-the-art algorithms on multiple datasets.The code has been released on https://github.com/houqianwen/MFFA-Net.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001214773000001  
Indexed Date:  
2024-05-11  
Institution Address:  
Qufu Normal Univ, Sch Cyber Sci & Engn, 57 Jingxuan West Rd, Jining 273165, Peoples R China  
Institution:

### 71. Deformable Convolution-Guided Multiscale Feature Learning and Fusion for

摘要: Object detection (OD) in unmanned aerial vehicle (UAV) images faces many challenges, with diverse-scale objects and small objects being particularly prominent issues. To alleviate these challenges, we propose a novel multiscale feature learning and feature fusion network under the guidance of deformable convolution. First, a deformable convolution-guided feature learning (DCGFL) block is designed in the backbone to extract more effective multiscale features. The DCGFL block leverages the adaptability of deformable convolution to the shapes and scales of objects, akin to spatial attention. Moreover, it also employs channel attention to identify important feature maps. Hence, the proposed backbone possesses the functionality of spatial attention and channel attention. Second, in the neck, we devise a simple generalized feature pyramid network (SimpleGFPN) with several deformable convolution-guided feature fusion (DCGFF) blocks to fuse multiscale features. The proposed neck has cross-layer and cross-scale pathways, facilitating effective information exchange and fusion between shallow spatial and deep semantic features. Third, the Scylla-IoU (SIoU) loss is used to better model the bounding box regression loss. Finally, the experimental results on the VisDrone2021 and UAVDT datasets show that the proposed method outperforms the compared OD methods. In terms of mean average precision, we obtain 37.8% on VisDrone2021 and 18.5% on UAVDT.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001164283100017  
Indexed Date:  
2024-03-24  
Institution Address:  
Xian Univ Architecture & Technol, Sch Informat & Control Engn, Xian 710055, Peoples R China  
Institution Address:  
Smart City Res Inst CETC, Shenzhen 518038, Peoples R China  
Institution:  
Institution:  
Smart City Res Inst CETC

### 72. Improving the Accuracy of Robot Collecting Organisms in Marine

摘要: In this paper, an improved YOLOv5 multiscale marine organism target detection algorithm (YOLOv5-Mult) is proposed to address the insufficient feature extraction ability of small targets, low detection accuracy, and high catching error of existing models in complex environments. First, real frame clustering is performed using the Kmeans++ method. Second, the BiFPN network module is adopted in lieu of the PANet network module to enhance the feature fusion ability. Next, the multilayer semantic fusion module RBC (RepBlock CSP) replaces the C3 module before the SPP layer of the Backbone network and the C3 module in the Neck layer to enrich the image semantic information. Finally, the multiscale feature fusion module MC (Mult Conv) replaces the last C3 module in the Backbone network to mitigate the semantic gap between different feature channel layers. Experimental results demonstrate that the improved algorithm attains a mAP value of 71.18%, which is 5.22% higher than that of the original YOLOv5 algorithm, providing accurate identification and fishing for underwater robots.  
An improved multi-scale marine organism target detection algorithm based on YOLOv5 (YOLOv5-Mult) has a higher mAP value than the original YOLOv5 algorithm by 5.22%. This algorithm enhances the target detection accuracy of underwater robots and effectively guides them in collecting organisms in the marine environment.image  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2513-0390  
入藏号:  
WOS:001112972800001  
Indexed Date:  
2023-12-16  
Institution Address:  
Northwest Normal Univ, Lanzhou 730070, Peoples R China  
Institution:

### 73. Image Super-Resolution Reconstruction Based on Partial Separation and

摘要: Currently,deep-learning-based super-resolution reconstruction networks suffer from issues such as convolution operation redundancy,incomplete image reconstruction information,and large model parameters that limit their applicability to edge devices.To address these issues,this study proposes a lightweight image superresolution reconstruction network based on partial separation and multiscale fusion.This network utilizes partial convolutions for feature extraction and separates partial image channels to reduce redundant computations while maintaining the quality of the image reconstruction.At the same time,a multiscale feature fusion module is designed to learn long-range dependency features and capture spatial features in the spatial dimension using a channel attention enhancement group.This reduces the loss of image reconstruction information and effectively restores the details and textures of the image.Finally,because the multiscale feature fusion block focuses on global feature extraction and fusion,an efficient inverted residual block is constructed to supplement the ability to extract local contextual information.The network is tested on five benchmark datasets:Set 5,Set 14,B 100,Urban 100,and Manga 109,with scale factors of 2,3,and 4 times.The parameters of the network are 373 000,382 0000,and 394 000,and the FLOPs are 84.0\*10~9,38.1\*10~9,and 22.1 \*10~9,respectively.Quantitative and qualitative experimental results show that compared with networks such as VDSR,IMDN,RFDN,and RLFN,the proposed network ensures image reconstruction quality with fewer network parameters.  
摘要:  
目前基于深度学习的超分辨率重建网络存在卷积运算冗余、图像重建信息不完整、模型参数庞大等问题,限制了其在边缘设备上的适用性。针对上述问题,提出一种轻量级的局部分离与多尺度融合图像超分辨率重建网络,该网络利用局部卷积对图像进行特征提取,通过分离部分图像通道,在减少网络冗余计算的同时保持图像重建的质量。设计一种多尺度特征融合模块,在空间维度学习长依赖特征,并采用一个通道注意力增强组在空间维度捕获空间特征,减少图像重建信息的丢失,有效恢复图像的细节纹理。由于多尺度特征融合模块更多地是从全局角度进行特征提取融合,因此构建一种高效反残差模块补充网络的局部上下文信息提取能力。在Set 5、Set 14、B 100、Urban 100、Manga 109这5个基准数据集上的实验结果表明,当尺度因子为2、3、4倍时,该网络的参数量分别为373 000、382 000、394 000,FLOPs分别为84.0\*10~9、38.1\*10~9、22.1\*10~9。与VDSR、IMDN、RFDN、RLFN等网络相比,该网络在较少网络参数的情况下,能够保证图像重建效果。  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1000-3428  
入藏号:  
CSCD:7752128  
Indexed Date:  
2024-08-30  
Institution Address:  
五邑大学智能制造学部, 江门, 广东 529020, 中国  
Institution Address:  
广东省科学院智能制造研究所;;华南理工大学机械与汽车工程学院, 广东省现代控制技术重点实验室;;, 广州;;广州, 广东;;广东 510070;;511442, 中国  
Institution Address:  
五邑大学智能制造学部;;广东省科学院智能制造研究所;;华南理工大学机械与汽车工程学院, ;;广东省现代控制技术重点实验室;;, 江门;;广州;;广州, 广东;;广东;;广东 529020;;510070;;511442, 中国  
Institution Address:  
广东省科学院智能制造研究所, 广东省现代控制技术重点实验室, 广州, 广东 510070, 中国  
Institution Address:  
School of Faculty of Intelligent Manufacturing,Wuyi University, Jiangmen, Guangdong 529020, China  
Institution Address:  
Institute of Intelligent Manufacturing,Guangdong Academy of Sciences;;School of Mechanical & Automotive Engineering,South China University of Technology, Guangdong Key Laboratory of Modern Control Technology;;, Guangzhou;;Guangzhou, Guangdong;;Guangdong 510070;;511442  
Institution Address:  
School of Faculty of Intelligent Manufacturing,Wuyi University;;Institute of Intelligent Manufacturing,Guangdong Academy of Sciences;;School of Mechanical & Automotive Engineering,South China University of Technology, ;;Guangdong Key Laboratory of Modern Control Technology;;, Jiangmen;;Guangzhou;;Guangzhou, Guangdong;;Guangdong;;Guangdong 529020;;510070;;511442  
Institution Address:  
Institute of Intelligent Manufacturing,Guangdong Academy of Sciences, Guangdong Key Laboratory of Modern Control Technology, Guangzhou, Guangdong 510070, China  
Institution:  
五邑大学智能制造学部  
Institution:  
广东省科学院智能制造研究所;;华南理工大学机械与汽车工程学院  
Institution:  
五邑大学智能制造学部;;广东省科学院智能制造研究所;;华南理工大学机械与汽车工程学院  
Institution:  
广东省科学院智能制造研究所  
Institution:  
School of Faculty of Intelligent Manufacturing,Wuyi University  
Institution:  
Institute of Intelligent Manufacturing,Guangdong Academy of Sciences;;School of Mechanical & Automotive Engineering,South China University of Technology  
Institution:  
School of Faculty of Intelligent Manufacturing,Wuyi University;;Institute of Intelligent Manufacturing,Guangdong Academy of Sciences;;School of Mechanical & Automotive Engineering,South China University of Technology  
Institution:  
Institute of Intelligent Manufacturing,Guangdong Academy of Sciences

### 74. Image Deblurring Based on Enhanced Multiscale Feature Network

摘要: Recently, multiscale and multistage image deblurring methods have encountered issues such as insufficient multiscale image feature extraction and loss of feature information due to stage deepening. To address the above problems, an image deblurring method based on an enhanced multiscale feature network is proposed in this paper. First, a multiscale residual feature extraction module is proposed, and convolution kernels with various sizes are used in the two branches to expand the receptive field and fully extract the feature information of images with various resolution sizes. Second, a crossstage attention module is proposed to filter and transfer the key features of the image. Finally, a cross-stage feature fusion module, similar to a jump connection, is designed to compensate for feature loss and fuse feature information from input images with various sizes, to enrich spatial feature information, and to improve texture processing. Experimental results on the GoPro and HIDE datasets show that the proposed method can successfully reconstruct the image.  
摘要:  
现有的多尺度多阶段图像去模糊方法存在多尺度图像特征提取利用不充分和阶段加深导致特征信息丢失的问题。针对上述问题,提出了一种基于增强多尺度特征网络的图像去模糊方法。首先,提出一个多尺度残差特征提取模块,在两个分支中使用不同大小的卷积核扩大感受野,充分提取不同分辨率大小图像的特征信息。其次,提出跨阶段注意力模块,过滤和传递图像的关键特征。最后,设计类似于跳跃连接的跨阶段特征融合模块,弥补特征损失,同时融合来自不同大小输入图像的特征信息,丰富空间特征信息,以更好地处理纹理细节。在GoPro和HIDE数据集上的实验结果表明,所提方法能够很好地重建图像。  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7370112  
Indexed Date:  
2023-04-22  
Institution Address:  
青岛大学计算机科学技术学院, 青岛, 山东 266071, 中国  
Institution Address:  
College of Computer Science & Technology,Qingdao University, Qingdao, Shandong 266071, China  
Institution:  
青岛大学计算机科学技术学院  
Institution:  
College of Computer Science & Technology,Qingdao University

### 75. LBP-based multi-scale feature fusion enhanced dehazing networks

摘要: Image dehazing is a prior process to perform advanced computer vision tasks such as certain target detection, and the degree of haze residue directly determines the performance of these tasks. Most existing dehazing methods follow a physical model of haze formation and obtain clear images indirectly by estimating global atmospheric light and transmission maps, but methods that rely on this model alone are difficult to use in real-world, complex hazy weather environments. In this paper, we propose an LBP-based multiscale feature fusion network for single-image dehazing to generate clear images directly end-to-end, and a multiscale feature fusion module is designed through an error feedback mechanism to alleviate a large amount of missing key feature information caused by the downsampling operation. A feature enhancement module using a Strengthen-Operate-subtract enhancement strategy is introduced into the decoder to improve the quality of the output image by refining the features of the image to be enhanced with the previously estimated image. The Inception module is added to the skip connection to alleviate the problem of excessive semantic gap in feature information at its two ends by increasing and deepening the width and depth of the network, and a self-attention mechanism is introduced to assign higher weights to key features. The above strategies enable the network to recover the haze images not only on the physical model of the input image; but also deep into the feature space to capture the internal correlation between each pixel. In addition, an improved LBP module is used to help the network obtain clearer detailed information and texture images through channel-by-channel matching of LBP images. Our model achieves a PSNR of 35.15 and an SSIM metric of 0.9905 on the SOTS dataset, and also performs well compared to state-of-the-art methods on images with different haze concentrations.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001040925900005  
Indexed Date:  
2023-08-23  
Institution Address:  
Shandong Technol & Business Univ, Coinnovat Ctr Shandong Collages & Univ Future Inte, Sch Informat & Elect Engn, Yantai 264005, Peoples R China  
Institution:

### 76. Shrimp larvae detection and counting in aquaculture using multiscale

摘要: Accurate detection and counting of shrimp larvae are essential for effective management and optimization of shrimp aquaculture systems. However, existing methods often fail in high-density environments, where larvae exhibit significant size variations and overlap, leading to high error rates and reduced applicability in real-world scenarios. To address these challenges, we propose a novel framework that integrates size-adaptive density map generation with a multiscale feature fusion network (MFFN), specifically designed for the dense and complex conditions of aquaculture. The proposed approach dynamically adjusts density map generation based on the varying sizes of shrimp larvae within an image, improving precision and enhancing detection accuracy. Simultaneously, the MFFN architecture extracts critical features across multiple scales, enabling the model to accurately detect and count larvae even in crowded scenes with significant occlusion. By combining these techniques, the framework demonstrates superior adaptability and robustness across diverse aquaculture environments. Extensive evaluations on a comprehensive shrimp larvae dataset show that our method achieves a detection accuracy of 93.68%, significantly outperforming traditional approaches. The model also reduces error rates and improves precision and recall metrics, with visual results showcasing its ability to distinguish and accurately count larvae of varying sizes. This work provides a valuable solution for automated shrimp larvae monitoring, supporting sustainable aquaculture practices. Future research will explore broader applications of the model and further refinements for enhanced scalability.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:001432683200001  
Indexed Date:  
2025-03-05  
Institution Address:  
Univ North Texas, Dept Elect Engn, Denton, TX 76205 USA  
Institution Address:  
Hanoi Univ Sci & Technol, Dept Elect Engn, Hanoi 112456, Vietnam  
Institution Address:  
Pukyong Natl Univ, Dept Artificial Intelligence & Convergence, Busan 48513, South Korea  
Institution:  
Institution:  
Institution:

### 77. A Lightweight Multiscale Feature Fusion Network for Solar Cell Defect

摘要: Solar cell defect detection is crucial for quality inspection in photovoltaic power generation modules. In the production process, defect samples occur infrequently and exhibit random shapes and sizes, which makes it challenging to collect defective samples. Additionally, the complex surface background of polysilicon cell wafers complicates the accurate identification and localization of defective regions. This paper proposes a novel Lightweight Multiscale Feature Fusion network (LMFF) to address these challenges. The network comprises a feature extraction network, a multi-scale feature fusion module (MFF), and a segmentation network. Specifically, a feature extraction network is proposed to obtain multi-scale feature outputs, and a multi-scale feature fusion module (MFF) is used to fuse multi-scale feature information effectively. In order to capture finer-grained multi-scale information from the fusion features, we propose a multi-scale attention module (MSA) in the segmentation network to enhance the network's ability for small target detection. Moreover, depthwise separable convolutions are introduced to construct depthwise separable residual blocks (DSR) to reduce the model's parameter number. Finally, to validate the proposed method's defect segmentation and localization performance, we constructed three solar cell defect detection datasets: SolarCells, SolarCells-S, and PVEL-S. SolarCells and SolarCells-S are monocrystalline silicon datasets, and PVEL-S is a polycrystalline silicon dataset. Experimental results show that the IOU of our method on these three datasets can reach 68.5%, 51.0%, and 92.7%, respectively, and the F1-Score can reach 81.3%, 67.5%, and 96.2%, respectively, which surpasses other commonly used methods and verifies the effectiveness of our LMFF network.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001397771900001  
Indexed Date:  
2025-01-23  
Institution Address:  
Guizhou Univ, Sch Mech Engn, Guiyang 550025, Peoples R China  
Institution Address:  
Beijing Jiaotong Univ, Sch Comp Sci & Technol, Beijing 100044, Peoples R China  
Institution:  
Institution:

### 78. Airport Near-Altitude Flying Birds Detection Based on Information

摘要: Airport near-altitude flying birds' detection is a critical task to avoid collisions of bird-to-aircraft. However, there still are two challenges as follows: 1) the inconspicuous features of birds make the detection task challenging, due to the small sizes and 2) deep networks are unfriendly to edge devices, while lightweight networks may exacerbate information loss due to depth separable convolution (DSC). In this article, we propose a lightweight network, named as ICFF-YOLOv5, to address the aforementioned problems and detect near-altitude flying birds effectively. To avoid the feature loss, we design a module of feature fusion, named as information compensation feature fusion (ICFF) to reduce the semantic difference between feature maps of different scales. Furthermore, we develop a double combination convolution (DCC) to reduce the number of parameters and achieve an advanced performance. In addition, we construct a dataset to evaluate the performance of our method. The experimental results demonstrate that the propose method can accurately detect the flying birds with varying shapes and scenarios, and achieve the values of evaluation metric mAP0.5 with 82.3 and F1 score with 82.4.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001087769200079  
Indexed Date:  
2023-11-24  
Institution Address:  
Tianjin Univ Technol, Res Ctr Learning Based Intelligent Syst, Key Lab Comp Vis & Syst, Minist Educ, Tianjin 300384, Peoples R China  
Institution Address:  
Tianjin Univ Technol, Sch Comp Sci & Engn, Tianjin 300384, Peoples R China  
Institution:  
Institution:

### 79. Multiscale Context-Aware Feature Fusion Network for Land-Cover

摘要: Recently, several land-cover classification models have achieved great success in terms of both accuracy and computational performance. However, it remains challenging due to interclass similarities, intraclass variations, scale-related inaccuracies, and high computational complexity. First, these methods fail to establish a correlation among different feature maps during multiscale feature extraction, leading to interclass similarities and intraclass variations. Second, they underutilize feature interdependencies of the context contained in each layer of the encoder-decoder architecture, causing scale-related inaccuracies. Third, they cause checkerboard artifacts and blurry edges, which can negatively impact the accuracy and generate segmentation map at increased computational cost. To address these problems, this article proposes a novel multiscale context-aware feature fusion network (MCN) for high-resolution urban scene images. MCN mainly consists of three modules: First, a multiscale feature enhancement module for backbone network to extract rich spatial information dynamically by incorporating dense correlation among feature maps with different receptive fields; second, multilayer feature fusion module as skip connections to produce a single high-level representation of the local-global context by capturing low-, mid-, and high-level interdependencies at different encoder-decoder stages; and third, pixel-shuffle decoder to reduce the blurry edges and checkerboard artifacts while upsampling with reduced number of parameters. Experiments on three high-resolution aerial and satellite urban scene datasets show that MCN consistently outperforms the mainstream land-cover classification models. Specifically, MCN achieves an OA of 93.51 on Potsdam, 90.18 on Vaihingen, and an mIoU of 73.73 on DeepGlobe.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001070732700017  
Indexed Date:  
2023-10-12  
Institution Address:  
Chongqing Univ, Sch Microelect & Commun Engn, Chongqing 400044, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Opt & Elect, Key Lab Opt Engn, Chengdu 610209, Peoples R China  
Institution:  
Institution:

### 80. Improved RetinaNet Algorithm for Object Detection

摘要: Based on the problems that the classical one-stage object detection algorithm RetinaNet is difficult to fully extract and fuse different stage features,while the bounding box regression is not sufficiently accurate,an improved RetinaNet algorithm for object detection is proposed.First,the algorithm adds multispectral channel attention to the feature extraction module,which incorporates more frequency components in the input features into the attention processing to capture the original rich information of the features.Thereafter,the multiscale feature fusion module is added after the feature extraction module,and the multiscale feature fusion module includes a path aggregation module and a feature fusion operation.The pathaggregation module enhances the information flow of the entire feature pyramid by building bottom-up paths and using accurate positioning signals on shallower feature layers.The feature fusion operation further enhances the fusion effect of multistage features by fusing the feature information from each stage.Finally,the Complete Intersection over Union(CIoU) loss function is introduced in the bounding box regression process.The loss function starts from three important geometric factors,namely,the overlapping area of the bounding box,the distance between the center points,and the aspect ratio to improve the convergence speed of the regression process and accuracy.The experimental results on the MS COCO and PASCAL VOC datasets show that,compared with the RetinaNet algorithm,the average accuracy of the improved RetinaNet algorithm on the two datasets is increased by 2.1 and 1.1 percentage points,especially for the MS COCO data set.For the detection of large targets,improving the detection accuracy is more significant.  
摘要:  
针对经典一阶段目标检测算法RetinaNet难以充分提取不同阶段特征、边界框回归不够准确等问题,提出一个面向目标检测的改进型RetinaNet算法。在特征提取模块中加入多光谱通道注意力,将输入特征中的频率分量合并到注意力处理中,从而捕获特征原有的丰富信息。将多尺度特征融合模块添加到特征提取模块,多尺度特征融合模块包括1个路径聚合模块和1个特征融合操作,路径聚合模块通过搭建自底向上的路径,利用较浅特征层上精确的定位信号增强整个特征金字塔的信息流,特征融合操作通过融合来自每个阶段的特征信息优化多阶段特征的融合效果。此外,在边界框回归过程中引入完全交并比损失函数,从边界框的重叠面积、中心点距离和长宽比这3个重要的几何因素出发,提升回归过程的收敛速度与准确性。在MS COCO数据集和PASCAL VOC数据集上的实验结果表明,与RetinaNet算法相比,改进型RetinaNet算法在2个数据集上的平均精度分别提高了2.1、1.1个百分点,尤其对于MS COCO数据集中较大目标的检测,检测精度的提升效果更加显著。  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
2  
ISSN:  
1000-3428  
入藏号:  
CSCD:7272665  
Indexed Date:  
2022-10-28  
Institution Address:  
郑州大学软件学院;;战略支援部队信息工程大学信息系统工程学院, ;;, 郑州;;郑州, ;; 450000;;450000  
Institution Address:  
战略支援部队信息工程大学信息系统工程学院, 郑州, 河南 450000, 中国  
Institution Address:  
School of Software, Zhengzhou University;;School of Information Systems Engineering, Strategic Support Force Information Engineering University, ;;, Zhengzhou;;Zhengzhou, ;; 450000;;450000  
Institution Address:  
School of Information Systems Engineering, Strategic Support Force Information Engineering University, Zhengzhou, Henan 450000, China  
Institution:  
郑州大学软件学院;;战略支援部队信息工程大学信息系统工程学院  
Institution:  
战略支援部队信息工程大学信息系统工程学院  
Institution:  
School of Software, Zhengzhou University;;School of Information Systems Engineering, Strategic Support Force Information Engineering University  
Institution:  
School of Information Systems Engineering, Strategic Support Force Information Engineering University

### 81. Generation of high-dimensional vibration signal and its application in

摘要: Imperfect data, such as data scarcity and imbalance, have a negative impact on intelligent fault diagnosis. Generative adversarial networks (GANs) have proven to be a potential method for augmenting data. However, the overall quality of data generated by GANs still requires improvement. In this study, a multiscale feature-fusion GAN was proposed for generating high-dimensional signals. A strategy that generates signals with dimensionality higher than the desired dimensionality was proposed. The generated signals were resampled to improve the phase diversity of the generated signals. To enhance the feature extraction capability of the GAN for high-dimensional signals, a multiscale feature extraction structure was designed. The integration of multiscale feature extraction and fusion was achieved without a significant increase in the computational burden. To enhance the amplitude diversity of the generated signals, a reconstruction network was designed that directly constrained the spatial distribution of the generated signals. Experimental results show that the model has advantages in terms of the similarity and diversity of the generated signals. The effectiveness of the model in fault diagnosis was verified using two motor datasets, where the fault diagnosis model using the sample set augmented by the proposed model obtained a 36.32% improvement in accuracy compared with using the original sample set. In addition, in comparison experiments, the model achieved a higher diagnostic accuracy improvement of 28.87% compared to four other published models.  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001072187600001  
Indexed Date:  
2023-10-06  
Institution Address:  
Xi An Jiao Tong Univ, Key Lab Educ, Minist Modern Design & Rotor Bearing Syst, Xian 710049, Peoples R China  
Institution Address:  
Univ British Columbia, Sch Engn, Vancouver, BC V6T 1Z4, Canada  
Institution:  
Institution:

### 82. Evaluation of Students' Innovation and Entrepreneurship Ability Based on

摘要: As the country's high-quality talents, college students are an important force in national construction. Evaluating the innovative and entrepreneurial abilities for Chinese students will help promote innovation and entrepreneurship education system and improve the reform of educational system and mechanism of colleges, thereby enhancing the innovation and entrepreneurship abilities of college students and then pushing the country into the ranks of a strong country in human resources and a strong country in talents. This work designs a ResNet-based evaluation method to college innovation and entrepreneurship abilities; the main contributions are as follows. (1) When ResNet performs feature extraction, there are problems of bloated network structure and feature loss. A feature extraction backbone network based on ResNet is proposed. To solve the issue of loss for shallow features in process of feature extraction, a skip architecture is added to fuse the shallow details and spatial information with the deep semantic information. To solve the problem of weak model generalization ability caused by the shallow network, a network stacking strategy is proposed to deepen the network structure. (2) Aiming at the problem that ResNet using single-scale feature prediction cannot effectively utilize multiscale features in the network, a multiscale feature prediction is designed. According to idea of feature pyramid, multiple feature maps with different scales are selected for the improved residual network. It designed a multiscale feature fusion strategy for fusing the selected multiscale feature maps into a feature map and evaluated the innovation and entrepreneurship abilities on the fused feature maps. Finally, comparative experiment proves that the improved feature extraction backbone network and multiscale feature scheme can improve performance accuracy on constructed dataset.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1574-017X  
eISSN:  
1875-905X  
入藏号:  
WOS:000772994100003  
Indexed Date:  
2022-04-14  
Institution Address:  
Hangzhou Vocat & Tech Coll, Human Resources Off, Hangzhou, Zhejiang, Peoples R China  
Institution Address:  
Hangzhou Vocat & Tech Coll, Deans Off, Hangzhou, Zhejiang, Peoples R China  
Institution:  
Institution:

### 83. BPR-Net: Balancing Precision and Recall for Infrared Small Target

摘要: Most current infrared small target detection methods attempt to fuse local and global information by using single-scale inputs and creating a multiscale feature pyramid during network feeding forward. Further to this, our research finds that using high-resolution inputs can improve recall, while low-resolution inputs improve precision. Nevertheless, solely focusing on global or local information can result in missing targets and false alarms. To address these issues, we propose the BPR-Net to balance precision and recall via a novel multiscale attention mechanism, which combines semantic and shallow features of multiscale inputs (MS). We first scale the input image into multiple images with varying resolutions and feed them into the network. In the encoder, the scale fusion module (SFM) fuses features from corresponding images of different resolutions. In the decoder, a channel fusion module (CFM) fuses useful information from multiple channels. Furthermore, a wavelet transform cross-layer skip layer (WTL) is employed to enhance the interaction between decoder layers for more effective multiscale feature fusion. Experimental results demonstrate that our approach achieves a balance between recall and precision and yields state-of-the-art performance on challenging benchmarks including Sirst, miss detection versus false alarm (MDvsFA), and small infrared aerial target detection (SIATD). Notably, our approach achieves an F1 score of 0.9409 on the challenging benchmark SIATD, surpassing the state-of-the-art method by 16.7%.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
9  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001043257600033  
Indexed Date:  
2023-08-24  
Institution Address:  
Huazhong Univ Sci & Technol, Key Lab Image Proc & Intelligent Control, Minist Educ, Wuhan 430074, Peoples R China  
Institution Address:  
Huazhong Univ Sci & Technol, Sch Artificial Intelligence & Automat, Minist Educ, Wuhan 430074, Peoples R China  
Institution:  
Institution:

### 84. A Multiscale Feature Pyramid SAR Ship Detection Network With Robust

摘要: Synthetic aperture radar (SAR) ship detection is widely used in cutting-edge applications such as environmental protection, traffic monitoring, search, and rescue. Lightweight detection algorithms are more important for practical applications. Although there has been extensive research in this field, there are some problems with the existing lightweight algorithms. For example, it is easy to misjudge targets that are mixed with the background, and the detection effect is not ideal for targets with few samples in the dataset. The root cause of these problems lies in the fact that the useless information in the background is relatively close to the target, and existing algorithms are too simplistic in fusing features at different levels, resulting in algorithms not being robust enough when facing these problems. Therefore, this article proposes a multiscale feature pyramid network (FPN)-based detection network (MFPNet), which introduces a spatial information-focusing module in the feature fusion channel to enhance the target's features to suppress interference information in the background and reduce misjudgment. Then, optimize the FPN and extract the importance of different resolution features based on network contribution to identifying multiscale targets. Experiments have shown that the MFPNet has better detection performance compared to existing algorithms on public datasets.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
被引频次合计:  
11  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001098810400009  
Indexed Date:  
2023-12-04  
Institution Address:  
Hunan Normal Univ, Key Lab Big Data Res & Applicat Basic Educ, Changsha 410081, Peoples R China  
Institution Address:  
Hunan Normal Univ, Coll Educ Sci, Changsha 410081, Peoples R China  
Institution Address:  
Univ Leicester, Sch Comp & Math Sci, Leicester LE1 7RH, England  
Institution:  
Institution:  
Institution:

### 85. LMFFNet: A Well-Balanced Lightweight Network for Fast and Accurate

摘要: Real-time semantic segmentation is widely used in autonomous driving and robotics. Most previous networks achieved great accuracy based on a complicated model involving mass computing. The existing lightweight networks generally reduce the parameter sizes by sacrificing the segmentation accuracy. It is critical to balance the parameters and accuracy for real-time semantic segmentation. In this article, we propose a lightweight multiscale-feature-fusion network (LMFFNet) mainly composed of three types of components: split-extract-merge bottleneck (SEM-B) block, feature fusion module (FFM), and multiscale attention decoder (MAD), where the SEM-B block extracts sufficient features with fewer parameters. FFMs fuse multiscale semantic features to effectively improve the segmentation accuracy and the MAD well recovers the details of the input images through the attention mechanism. Without pretraining, LMFFNet-3-8 achieves 75.1% mean intersection over union (mIoU) with 1.4 M parameters at 118.9 frames/s using RTX 3090 GPU. More experiments are investigated extensively on various resolutions on other three datasets of CamVid, KITTI, and WildDash2. The experiments verify that the proposed LMFFNet model makes a decent tradeoff between segmentation accuracy and inference speed for real-time tasks. The source code is publicly available at https://github.com/Greak-1124/LMFFNet.  
在 BIOSIS Citation Index 中的被引频次:  
5  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
64  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
66  
ISSN:  
2162-237X  
eISSN:  
2162-2388  
入藏号:  
WOS:000805797800001  
PubMed ID:  
35622806  
Indexed Date:  
2022-06-14  
Institution Address:  
Jinan Univ, Coll Informat Sci & Technol, Dept Elect Engn, Guangzhou 510632, Peoples R China  
Institution Address:  
Chinese Acad Sci, Shanghai Adv Res Inst, Shanghai 201210, Peoples R China  
Institution Address:  
Univ Macau, Dept Comp & Informat Sci, Macau 999078, Peoples R China  
Institution:  
Institution:  
Institution:

### 86. MFCNet: Multimodal Feature Fusion Network for RGB-T Vehicle Density

摘要: The basic task of vehicle density estimation is to use image information to estimate the distribution and quantity of vehicles within it. However, many previous methods only use the optical information in red-green-blue (RGB) images, which makes it difficult to effectively identify potential vehicles under poor light, strong reflections, and bad weather, resulting in unsatisfactory density estimation performance. To address these problems, we consider introducing thermal images to provide a richer source of information for the vehicle density estimation task, and propose a multimodal feature fusion network (MFCNet) for accurate RGB-Thermal (RGB-T) vehicle density estimation. First, multimodal features are cross-integrated through the attention-guided multiscale feature fusion coordination module (MFFC) to compensate for the limitations of single modal features. Following this, the edge feature calibration module (EFC) is utilized to correct the spatial misalignment regions between modalities. Subsequently, the adaptive deep fusion module (ADFM) is applied to further refine the features on the global scale and improve the intermodality correlation. Finally, the features of different stages are fused step by step to obtain the final fused feature, which is fed into a simple regression header to generate a pixel-level vehicle density map. Experimental results show that the GAME2 and root mean square error of the proposed method are reduced to 5.21 and 3.54 on the DroneVehicle dataset, respectively. Compared with existing vehicle density estimation methods, MFCNet achieves competitive accuracy and can be applied to the vehicle density estimation task in unconstrained scenarios. Our codes will be available at https://github.com/QLingX/MFCNet.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2327-4662  
入藏号:  
WOS:001414873100005  
Indexed Date:  
2025-02-14  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao 266590, Peoples R China  
Institution:

### 87. Dynamic Convolution Covariance Network Using Multiscale Feature Fusion

摘要: The rapid increase in spatial resolution of remote sensing scene images (RSIs) has led to a concomitant increase in the complexity of the spatial contextual information contained therein. The coexistence of numerous smaller features makes it challenging to accurately locate and mine these features, which in turn makes accurate interpretation difficult. In order to address the aforementioned issues, this article proposes a dynamic convolution covariance network (ODFMN) based on omni-dimensional dynamic convolution, which can extract multidimensional and multiscale features from RSIs and perform statistical higher-order representation of feature information. First, in order to fully exploit the complex spatial context information of RSIs and at the same time improve the limitation of a single static convolution kernel for feature extraction, we constructed a omni-dimensional feature extraction module based on dynamic convolution, which fully extracts the 4-D information within the convolution kernel. Then, to make full use of the full-dimensional feature information extracted from each level in the network, the feature representation is enriched by constructing multiscale feature fusion module to establish relationships from local to global. Finally, higher order statistical information is employed to address the challenge of representing first-order information for smaller object features, which is inherently difficult to do. Experiments conducted on publicly available datasets have demonstrated that the method achieves high classification accuracies of 99.04%, 95.34%, and 92.50%, respectively. Furthermore, the method has been verified to have high capture accuracy for feature target contours, shapes, and spatial context information through feature visualization.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001322151100006  
Indexed Date:  
2024-10-10  
Institution Address:  
Tianjin Univ Technol, Sch Comp Sci & Engn, Tianjin 300384, Peoples R China  
Institution Address:  
Tianjin Univ Technol, Key Lab Comp Vis & Syst, Minist Educ, Tianjin 300384, Peoples R China  
Institution:  
Institution:

### 88. Image Deblurring Based on Enhanced Multiscale Feature Network

摘要: Recently, multiscale and multistage image dehlurring methods have encountered issues such as insufficient multiscale image feature extraction and loss of feature information due to stage deepening. To address the above problems, an image dehlurring method based on an enhanced multiscale feature network is proposed in this paper. First, a multiscale residual feature extraction module is proposed, and convolution kernels with various sizes are used in the two branches to expand the receptive field and fully extract the feature information of images with various resolution sizes. Second, a cross stage attention module is proposed to filter and transfer the key features of the image. Finally, a cross-stage feature fusion module, similar to a jump connection, is designed to compensate for feature loss and fuse feature information from input images with various sizes, to enrich spatial feature information, and to improve texture processing. Experimental results on the G oPro and HIDE datasets show that the proposed method can successfully reconstruct the image.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
1006-4125  
入藏号:  
WOS:000892702100027  
Indexed Date:  
2022-12-23  
Institution Address:  
Qingdao Univ, Coll Comp Sci & Technol, Qingdao 266071, Shandong, Peoples R China  
Institution:

### 89. Image Semantic Segmentation Algorithm with Multi-scale Feature Fusion

摘要: For the problems of small-scale target losing and discontinuous segmentation in existing image semantic segmentation,an image semantic segmentation algorithm with multi-scale feature fusion and enhancement is proposed.Based on DeeplabV3+ network,the algorithm improves the ability to describe small target features by building a multi-scale feature extraction and fusion enhancement network.The network can also obtain small target feature while segmenting large targets,so it can solve the problem of the small target losing and the discontinuous segmentation in the semantic segmentation.Experimental results on the Cityscapes dataset show that the improved algorithm significantly improves the accuracy of small target segmentation and optimizes the problem of discontinuous segmentation.Finally,the generalization of the improved algorithm is verified on the public dataset PASCAL VOC 2012.  
摘要:  
针对现有的图像语义分割算法存在小尺度目标丢失和分割不连续的问题,提出多尺度融合增强的图像语义分割算法,该算法在DeeplabV3+网络模型的基础上,通过构建多尺度特征提取和融合增强网络提高了对小目标特征的描述能力,使网络在分割大目标的同时也能获得小目标的特征信息,从而解决了语义分割时小尺度目标丢失和分割不连续的问题。在Cityscapes数据集上实验的结果表明,改进后的算法明显提升了小目标分割精度,解决了分割不连续的问题。最后在公开数据集PASCAL VOC 2012上进一步验证了改进算法的泛化性。  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1002-8331  
入藏号:  
CSCD:6924317  
Indexed Date:  
2021-06-04  
Institution Address:  
北京建筑大学电气与信息工程学院;;北京建筑大学, ;;建筑大数据智能处理方法研究北京市重点实验室, ;;, 北京;;北京 100044;;100044, 中国  
Institution Address:  
School of Electrical and Information Engineering,Beijing University of Civil Engineering and Architecture Beijing;;Beijing University of Civil Engineering and Architecture, ;;Beijing Key Laboratory of Intelligent Processing for Building Big Data, ;;, Beijing;;Beijing 100044;;100044  
Institution:  
北京建筑大学电气与信息工程学院;;北京建筑大学  
Institution:  
School of Electrical and Information Engineering,Beijing University of Civil Engineering and Architecture Beijing;;Beijing University of Civil Engineering and Architecture

### 90. Improved YOLOv4-Tiny Target Detection Method Based on Adaptive

摘要: To improve the accuracy of material identification under low contrast conditions, this paper proposes an improved YOLOv4-tiny target detection method based on an adaptive self-order piecewise enhancement and multiscale feature optimization. The model first constructs an adaptive self-rank piecewise enhancement algorithm to enhance low-contrast images and then considers the fast detection ability of the YOLOv4-tiny network. To make the detection network have a higher accuracy, this paper adds an SE channel attention mechanism and an SPP module to this lightweight backbone network to increase the receptive field of the model and enrich the expression ability of the feature map. The network can pay more attention to salient information, suppress edge information, and effectively improve the training accuracy of the model. At the same time, to better fuse the features of different scales, the FPN multiscale feature fusion structure is redesigned to strengthen the fusion of semantic information at all levels of the network, enhance the ability of network feature extraction, and improve the overall detection accuracy of the model. The experimental results show that compared with the mainstream network framework, the improved YOLOv4-tiny network in this paper effectively improves the running speed and target detection accuracy of the model, and its mAP index reaches 98.85%, achieving better detection results.  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
eISSN:  
2076-3417  
入藏号:  
WOS:001037875700001  
Indexed Date:  
2023-08-08  
Institution Address:  
Yanshan Univ, Sch Elect Engn, Qinhuangdao 066004, Peoples R China  
Institution Address:  
Guangxi LiuGong Machinery Co Ltd, Intelligent Technol Res Inst, Global Res & Dev Ctr, Liuzhou 545007, Peoples R China  
Institution Address:  
Univ Elect Sci & Technol, Sch Resources & Environm, Chengdu 611731, Peoples R China  
Institution Address:  
Guangxi Univ Sci & Technol, Sch Automat, Liuzhou 545006, Peoples R China  
Institution Address:  
Guangxi Univ Sci & Technol, Guangxi Collaborat Innovat Ctr Earthmoving Machine, Liuzhou 545006, Peoples R China  
Institution:  
Institution:  
Guangxi LiuGong Machinery Co Ltd  
Institution:  
Institution:  
Institution:

### 91. Multiscale Feature Adaptive Fusion for Object Detection in Optical

摘要: So far, the accuracy of object detection in natural images has been continuously improved, and at the same time, object detection has received more and more attention in the field of remote sensing. However, different from natural images, remote sensing images have a large number of multiscale objects and large, complex backgrounds. As the size of remote sensing images is generally large, small objects are easy to lose information after downsampling. To deal with these problems, we propose a multiscale feature adaptive fusion (MFAF) method. Specifically, we first use the multiscale feature integration (FI) module and the spatial attention weight (SAW) module to construct the feature fusion module to achieve the adaptive fusion of multiscale features. Then, we add a detail enhancement (DE) module at the back of the backbone to enhance the quality of features in each scale. Next, we obtain the relationship between the features of different channels by embedding a squeeze and excitation (SE) block to highlight the useful features. Finally, a cross stage partial (CSP) block is adopted to replace continuous convolutions to reduce the number of parameters and the loss of features. We implement our work in YOLOv4. In experiments, we evaluate our method on large-scale remote sensing image datasets HRRSD and DIOR, and, respectively, achieve improvements of 2.7% and 1.9% in AP(50) compared with YOLOv4, successfully improving the multiscale object detection performance.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
9  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000809407200001  
Indexed Date:  
2022-06-20  
Institution Address:  
Nanjing Normal Univ, Sch Elect & Automat Engn, Nanjing 210023, Peoples R China  
Institution Address:  
Nanjing Normal Univ, Jiangsu Open Lab Major Sci Instrument & Equipment, Nanjing 210023, Peoples R China  
Institution Address:  
Datang Jiangsu Power Generat Co Ltd, Nanjing 210011, Peoples R China  
Institution:  
Institution:  
Institution:  
Datang Jiangsu Power Generat Co Ltd

### 92. Study on Image Classification Algorithm Based on Multi-Scale Feature

摘要: This paper introduces the MMTADAN, an innovative algorithm designed to enhance cross-domain image classification. By integrating multi-scale feature extraction with Taylor series-based detail enhancement and adversarial domain adaptation, the MMTADAN effectively aligns features between the source and target domains. The proposed approach addresses the critical challenge of generalizing classification models across diverse datasets, demonstrating significant improvements in performance. The findings suggest that retaining essential image details through multi-scale extraction and Taylor series enhancement can lead to better classification outcomes, making the MMTADAN a valuable contribution to the field of image classification.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001366796000001  
Indexed Date:  
2024-12-07  
Institution Address:  
Dalian Maritime Univ, Marine Engn Coll, Dalian 116000, Peoples R China  
Institution Address:  
Univ Manchester, Dept Elect & Elect Engn, Manchester M13 9PL, England  
Institution Address:  
Univ Nottingham, Sch Math Sci, Nottingham NG9 2SE, England  
Institution Address:  
China Univ Geosci, Natl Engn Res Ctr Geog Informat Syst, Wuhan 430078, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 93. Single-image dehazing via a coarse-to-fine multiscale approach

摘要: A single-image dehazing method based on a coarse-to-fine, multiscale approach is proposed to address the shortcomings of existing dehazing methods, which often result in loss of detailed texture, color distortion, or incomplete processing of non-uniform dense fog. First, features are extracted from foggy images using the residual feature attention module of the backbone network. Second, images at different scales are preprocessed using convolution, and the preprocessed shallow features are fused with the backbone network through a multiscale feature fusion module. Asymmetric features at different levels are then effectively fused. Finally, shallow information is adaptively combined with deep information at the output. Compared with contrastive regularization loss, positive and negative sample information is constructed to make the defogged image more similar to a fog-free image. The experimental results show that, compared with existing representative defogging methods, the proposed method effectively defogs synthetic and real datasets and outperforms comparison methods in terms of detail retention and color reproduction.  
摘要:  
为了解决现有图像去雾算法易出现细节纹理丢失、颜色失真或对非均匀浓雾处理不彻底的问题,提出一种基于粗到细的多尺度单幅图像去雾方法。首先,主干网络使用残差特征注意力模块对有雾图像进行特征提取;其次,将不同尺度的输入图像进行卷积预处理,通过多尺度特征融合模块将预处理的浅层特征与主干网络融合;再次,将不同粒度的非对称特征进行有效融合;最后,将浅层信息与深层信息自适应混合输出,通过对比正则损失构建正负样本信息,使得去雾图像更接近无雾图像。实验结果表明,与已有代表性的去雾方法相比,提出的方法能对合成数据集与真实数据集进行有效去雾,在细节保留与色彩还原上优于对比方法。  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1673-4785  
入藏号:  
CSCD:7852586  
Indexed Date:  
2025-02-09  
Institution Address:  
(保定)华北电力大学控制与计算机工程学院;;(保定)华北电力大学, ;;复杂能源系统智能计算教育部工程研究中心, 保定;;保定, 河北;;河北 071003;;071003, 中国  
Institution Address:  
(保定)华北电力大学控制与计算机工程学院, 保定, 河北 071003, 中国  
Institution Address:  
School of Control and Computer Engineering,North China Electric Power University(Baoding);;North China Electric Power University(Baoding), ;;Engineering Research Center of the Ministry of Education for Intelligent Computing of Complex Energy System, Baoding;;Baoding, ;; 071003;;071003  
Institution Address:  
School of Control and Computer Engineering,North China Electric Power University(Baoding), Baoding, Hebei 071003, China  
Institution:  
(保定)华北电力大学控制与计算机工程学院;;(保定)华北电力大学  
Institution:  
(保定)华北电力大学控制与计算机工程学院  
Institution:  
School of Control and Computer Engineering,North China Electric Power University(Baoding);;North China Electric Power University(Baoding)  
Institution:  
School of Control and Computer Engineering,North China Electric Power University(Baoding)

### 94. RG-YOLO: multi-scale feature learning for underwater target detection

摘要: Underwater target detection is pivotal in advancing marine development. However, the intricate underwater environments pose substantial challenges to this task, including dense target distributions, low contrast, and limited lighting conditions. To improve the performance of underwater target detection, this paper proposes an underwater target detection model based on YOLOv7-tiny framework, named RG-YOLO. Firstly, an efficient reparameterized multi-scale fusion (RMF) module is designed to boost the feature extraction capability of the model, capturing both local and global information and learning richer multi-scale feature representations. Secondly, this paper proposes a gather and distribute feature pyramid network (GDFPN) to significantly enhance the ability of multi-scale feature fusion. In addition, a dynamic head module is incorporated into the network. This module elevates detection performance for underwater small targets by unifying object detection heads with attentions. Finally, Shape-IoU is introduced to emphasize the inherent attributes of the bounding boxes, thereby obtaining higher accuracy in localization. Experimental results on the UTDAC2020, DUO, and RUOD datasets demonstrate that RG-YOLO achieves mean average precision (mAP) @50 of 85.1%, 86.1%, and 85.7%, respectively. These results represent improvements of 2.9%, 1.6%, and 2.3% over the baseline model, showing superior performance in underwater target detection compared to other state-of-the-art models.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
0942-4962  
eISSN:  
1432-1882  
入藏号:  
WOS:001382307200003  
Indexed Date:  
2024-12-28  
Institution Address:  
Shanghai Maritime Univ, Coll Informat Engn, Shanghai 201306, Peoples R China  
Institution:

### 95. Confidence-Based Weakly Supervised for Aircraft Detection From Remote

摘要: Weakly supervised aircraft detection is a crucial topic in object detection. The existing approach uses a class activation map (CAM) to generate a pseudo-label. However, the reliability of the pseudo-label is often overlooked as current methods mainly focus on expanding the activating regions. Weak supervision relies on confidence to determine pseudo-labeling reliability. Pseudo-label confidence has often been measured by the peak values of the predicted probability distribution, which is essentially a bootstrap method that is not objective enough to provide precise predictions. In this letter, we propose a confidence-based weakly supervised framework (CWSF) to improve the reliability of pseudo-labels. A multiscale feature fusion and refinement module (MFFR) is designed to optimize and extract complex features in remote sensing images. We also introduce a confidence predictor (CP) to determine pseudo-label confidence. Moreover, the loss function (CWSLoss) is proposed to train the network framework. The proposed CWSF improves the quality of pseudo-labeling and object detection results. Our experiments on selected datasets NPWU VHR-10 (67.01%) and PatternNet (65.42%) show that CWSF significantly outperforms previous state-of-the-art methods.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001205790900002  
Indexed Date:  
2024-05-02  
Institution Address:  
Guangxi Univ, Sch Elect Engn, Nanning 530004, Peoples R China  
Institution:

### 96. Video action recognition with Key-detail Motion Capturing based on

摘要: At present, existing research works on action recognition are still not ideal, when most of the video content is redundant such as video clips without any object motion, and human actions in the video are complex. The reasons are as follows: (1) Most of them lack attention to key-motion information of the video, thus irrelevant information will be input into the model. (2) And there is a lack of interaction between video spatial and temporal information, which may cause the loss of detailed motion information in the video. In this paper, we propose a Key-detail Motion Capturing Network (K-MCN) to solve these problems, which contains two modules. The first one is the Video Key-motion Spectrum Analyzer (VKSA) module. In this module, the video optical flow can be subjected to frequency spectrum analysis, filtering and clustering to extract the key-motion frames. The second one is the Multiscale Motion Spatiotemporal Interaction module, which allows multi-scale modeling and fusion of spatial and temporal features extracted from key-motion frames, enabling the network to realize the interaction and supplement of multiscale spatiotemporal information. Finally, we conducted extensive experiments on the UCF101, HMDB51 and Something-SomethingV1 datasets, and the results showed that our method achieves better performance compared with other state-of-the-art methods.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:000741634600006  
Indexed Date:  
2022-01-19  
Institution Address:  
Guangdong Univ Technol, Sch Comp, Guangzhou 510006, Peoples R China  
Institution Address:  
Univ Macau, Dept Comp & Informat Sci, Taipa 999078, Macao, Peoples R China  
Institution Address:  
South China Univ Technol, Sch Comp Sci & Engn, Guangzhou 510006, Peoples R China  
Institution Address:  
Guangdong Univ Technol, Sch Informat Engn, Guangzhou 510006, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 97. Forecasting Variation Trends of Stocks via Multiscale Feature Fusion and

摘要: Forecasting stock price trends accurately appears a huge challenge because the environment of stock markets is extremely stochastic and complicated. This challenge persistently motivates us to seek reliable pathways to guide stock trading. While the Long Short-Term Memory (LSTM) network has the dedicated gate structure quite suitable for the prediction based on contextual features, we propose a novel LSTM-based model. Also, we devise a multiscale convolutional feature fusion mechanism for the model to extensively exploit the contextual relationships hidden in consecutive time steps. The significance of our designed scheme is twofold. (1) Benefiting from the gate structure designed for both long- and short-term memories, our model can use the given stock history data more adaptively than traditional models, which greatly guarantees the prediction performance in financial time series (FTS) scenarios and thus profits the prediction of stock trends. (2) The multiscale convolutional feature fusion mechanism can diversify the feature representation and more extensively capture the FTS feature essence than traditional models, which fairly facilitates the generalizability. Empirical studies conducted on three classic stock history data sets, i.e., S&P 500, DJIA, and VIX, demonstrated the effectiveness and stability superiority of the suggested method against a few state-of-the-art models using multiple validity indices. For example, our method achieved the highest average directional accuracy (around 0.71) on the three employed stock data sets.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1058-9244  
eISSN:  
1875-919X  
入藏号:  
WOS:000703631200002  
Indexed Date:  
2021-10-14  
Institution Address:  
Nanjing Forestry Univ, Coll Econ & Management, Nanjing 210037, Peoples R China  
Institution Address:  
China Shipbldg Ind Corp, China Ship Sci Res Ctr, Wuxi 214000, Jiangsu, Peoples R China  
Institution:  
Institution:

### 98. Insu-YOLO: An Insulator Defect Detection Algorithm Based on Multiscale

摘要: To keep the balance of precision and speed of unmanned aerial vehicles (UAVs) in detecting insulator defects during power inspection, an improved insulator defect identification algorithm, Insu-YOLO, which is based on the latest YOLOv8 network, is proposed in this paper. Firstly, to lower the computational complexity of the network, the GSConv module is introduced in the backbone and neck network. In the neck network, a lightweight content-aware reassembly of features (CARAFE) structure is adopted to better utilize the feature information for upsampling, which enhances the feature fusion capability of Insu-YOLO. Additionally, Insu-YOLO enhances the fusion between shallow and deep feature maps by adding an extra object detection layer, thereby increasing the accuracy for detecting small targets. The experimental results indicate that the mean average precision of Insu-YOLO reaches 95.9%, which is 3.95% higher than the YOLOv8n baseline model, with a memory usage of 9.2 MB. Moreover, the detection speed of Insu-YOLO is 87 frames/s which achieves the purpose of real-time identification of insulator defects.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
28  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
28  
ISSN:  
2079-9292  
入藏号:  
WOS:001045344300001  
Indexed Date:  
2023-08-19  
Institution Address:  
China Jiliang Univ, Sch Mech & Elect Engn, Hangzhou 310018, Peoples R China  
Institution:

### 99. Efficient Vehicle Detection in Remote Sensing Images with Bi-Directional

摘要: Facing with the challenges of the vehicle detection in remote sensing images,such as complex backgrounds,multi-scale differences,and difficulty in detecting small targets,a detection method GEM\_YOLO based on bidirectional multi-scale feature fusion is proposed.There are three main parts in this method:the first one is a globally efficient attention module that is designed as a feature extractor to achieve lightweight and efficient feature extraction,in order to solve the problem of object detection in complex backgrounds.Secondly,a bidirectional multi-scale feature fusion network is proposed as a feature fusion device,which adopts top-down and bottom-up feature fusion strategies to effectively promote information exchange between features at different levels.Finally,the application of an attention based on the dynamic detection head as a predictor enhances the perception of different scales,spatial positions,and tasks,further improving the accuracy and robustness of object detection.Related experiments are conducted on public datasets DIOR and DOTA,whose average accuracy reaches 92.4% and 81.4% that is significantly superior to other mainstream detection methods.Meanwhile,the fewer parameters and computational complexity provide an efficient solution for vehicle detection within the domain of remote sensing image detection.  
摘要:  
针对遥感图像中车辆检测面临的背景复杂、多尺度差异和小目标难以检测等挑战,提出了一种基于双向多尺度特征融合的检测方法GEM\_YOLO。该方法包括三个主要部分:设计了全局高效注意力模块作为特征提取器,实现轻量化和高效率的特征提取,以解决复杂背景下的目标检测问题;提出了双向多尺度特征融合网络作为特征融合器,采用自顶向下和自底向上的特征融合策略,有效促进不同层次特征之间的信息交互;应用基于注意力的动态检测头作为预测器,增强了对不同尺度、空间位置和任务的感知,进一步提升了目标检测的精度和鲁棒性。在公开数据集DIOR和DOTA上进行相关实验,该方法的平均精度均值达到92.4%和81.4%,显著优于其他主流检测方法,同时具有更少的参数量和计算量,为遥感图像检测领域中的车辆检测提供了一种高效解决方案。  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7753173  
Indexed Date:  
2024-08-30  
Institution Address:  
辽宁工程技术大学软件学院, 葫芦岛, 辽宁 125105, 中国  
Institution Address:  
School of Software,Liaoning Technical University, Huludao, Liaoning 125105, China  
Institution:  
辽宁工程技术大学软件学院  
Institution:  
School of Software,Liaoning Technical University

### 100. VMF-SSD: A Novel V-Space Based Multi-Scale Feature Fusion SSD for Apple

摘要: Apple leaf diseases seriously affect the quality of apples and may lead to yield losses, detecting apple leaf diseases accurately can prevent diseases from spreading and promote the healthy growth of the industry. However, recent studies cannot achieve accurate detection of leaf diseases with high accuracy because the lesions are of different sizes. So, this paper proposed a novel apple leaf disease detection method called VMF-SSD (V-space-based Multi-scale Feature-fusion SSD), which is designed to extract more reliable multi-scale feature representations for varied sizes of diseased spots and improve the final detection performance. The multi-scale feature extraction is established with multi-scale feature representation to further improve the disease detection performance, especially for small spots. After that, a V-space-based location branch is presented to enhance the texture feature information and help further identify disease spot location. Finally, attention mechanisms are utilized to automatically learn the importance of feature channels at different scales for distinguishing diseased spots of different sizes. Experimental results showed that the VMF-SSD method achieves 83.19% mAP and obtains the detection speed of 27.53 FPS on the test set, which indicates that the proposed VMF-SSD method can achieve competitive performance on apple leaf diseases detection task and satisfy the requirements of agricultural production applications.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
18  
在 BIOSIS Citation Index 中的被引频次:  
5  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
20  
ISSN:  
1545-5963  
eISSN:  
1557-9964  
入藏号:  
WOS:001006656100036  
PubMed ID:  
37015544  
Indexed Date:  
2023-07-26  
Institution Address:  
Northwest A&F Univ, Coll Informat Engn, Xianyang 712100, Peoples R China  
Institution Address:  
Minist Agr & Rural Affairs, Key Lab Agr Internet Things, Yangling 712100, Peoples R China  
Institution Address:  
Shaanxi Key Lab Agr Informat Percept & Intelligent, Yangling 712100, Peoples R China  
Institution:  
Institution:  
Institution:  
Shaanxi Key Lab Agr Informat Percept & Intelligent

### 101. ELCD: Efficient Lunar Crater Detection Based on Attention Mechanisms and

摘要: The detection and counting of lunar impact craters are crucial for the selection of detector landing sites and the estimation of the age of the Moon. However, traditional crater detection methods are based on machine learning and image processing technologies. These are inefficient for situations with different distributions, overlaps, and crater sizes, and most of them mainly focus on the accuracy of detection and ignore the efficiency. In this paper, we propose an efficient lunar crater detection (ELCD) algorithm based on a novel crater edge segmentation network (AFNet) to detect lunar craters from digital elevation model (DEM) data. First, in AFNet, a lightweight attention mechanism module is introduced to enhance the feature extract capabilities of networks, and a new multiscale feature fusion module is designed by fusing different multi-level feature maps to reduce the information loss of the output map. Then, considering the imbalance in the classification and the distributions of the crater data, an efficient crater edge segmentation loss function (CESL) is designed to improve the network optimization performance. Lastly, the crater positions are obtained from the network output map by the crater edge extraction (CEA) algorithm. The experiment was conducted on the PyTorch platform using two lunar crater catalogs to evaluate the ELCD. The experimental results show that ELCD has a superior detection accuracy and inference speed compared with other state-of-the-art crater detection algorithms. As with most crater detection models that use DEM data, some small craters may be considered to be noise that cannot be detected. The proposed algorithm can be used to improve the accuracy and speed of deep space probes in detecting candidate landing sites, and the discovery of new craters can increase the size of the original data set.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
被引频次合计:  
8  
eISSN:  
2072-4292  
入藏号:  
WOS:000875154600001  
Indexed Date:  
2022-11-07  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Comp Sci & Technol, Nanjing 211106, Peoples R China  
Institution:

### 102. MiniYOLO: A lightweight object detection algorithm that realizes the

摘要: The object detection task is to locate and classify objects in an image. The current state-of-the-art high-accuracy object detection algorithms rely on complex networks and high computational cost. These algorithms have high requirements on the memory resource and computing capability of the deployed device, and are difficult to apply to mobile and embedded devices. Through the depthwise separable convolution and multiple efficient network structures, this paper designs a lightweight backbone network and two different multiscale feature fusion structures, and proposes a lightweight one-stage object detection algorithm-MiniYOLO. With the model size of only 4.2 MB, MiniYOLO still maintains a high detection accuracy, realizing the trade-off between the model size and detection accuracy. Experimental results on MS COCO 2017 data set show that compared to the state-of-the-art PP-YOLO-tiny, MiniYOLO achieves higher mAP with the same model size. Compared with other lightweight object detection algorithms, MiniYOLO has certain advantages in detection accuracy or model size. The code associated with this paper can be downloaded from .  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
0884-8173  
eISSN:  
1098-111X  
入藏号:  
WOS:000856478200001  
Indexed Date:  
2022-09-28  
Institution Address:  
Northeastern Univ, Dept Software Engn, Shenyang, Peoples R China  
Institution Address:  
Northeastern Univ, Dept Comp Applicat Technol, Shenyang, Peoples R China  
Institution:  
Institution:

### 103. ReBiDet: An Enhanced Ship Detection Model Utilizing ReDet and

摘要: To enhance ship detection accuracy in the presence of complex scenes and significant variations in object scales, this study introduces three enhancements to ReDet, resulting in a more powerful ship detection model called rotation-equivariant bidirectional feature fusion detector (ReBiDet). Firstly, the feature pyramid network (FPN) structure in ReDet is substituted with a rotation-equivariant bidirectional feature fusion feature pyramid network (ReBiFPN) to effectively capture and enrich multiscale feature information. Secondly, K-means clustering is utilized to group the aspect ratios of ground truth boxes in the dataset and adjust the anchor size settings accordingly. Lastly, the difficult positive reinforcement learning (DPRL) sampler is employed instead of the random sampler to address the scale imbalance issue between objects and backgrounds in the dataset, enabling the model to prioritize challenging positive examples. Through numerous experiments conducted on the HRSC2016 and DOTA remote sensing image datasets, the effectiveness of the proposed improvements in handling complex environments and small object detection tasks is validated. The ReBiDet model demonstrates state-of-the-art performance in remote sensing object detection tasks. Compared to the ReDet model and other advanced models, our ReBiDet achieves mAP improvements of 3.20, 0.42, and 1.16 on HRSC2016, DOTA-v1.0, and DOTA-v1.5, respectively, with only a slight increase of 0.82 million computational parameters.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
5  
eISSN:  
2076-3417  
入藏号:  
WOS:001014078700001  
Indexed Date:  
2023-07-06  
Institution Address:  
Acad Mil Sci, Inst Syst Engn, Beijing 100000, Peoples R China  
Institution Address:  
Peking Univ, Sch Comp Sci, Beijing 100000, Peoples R China  
Institution:  
Acad Mil Sci  
Institution:

### 104. A Novel Adaptive Edge Aggregation and Multiscale Feature Interaction

摘要: Object detection (OD) in remote sensing (RS) images is an important task in the field of computer vision. OD techniques have achieved impressive advances in recent years. However, complex background interference, large-scale variations, and dense instances pose significant challenges for OD. These challenges may lead to misalignment between features extracted by OD models and the features of real objects. To address these challenges, we explore a novel single-stage detection framework for the adaptive fusion of multiscale features and propose a novel adaptive edge aggregation and multiscale feature interaction detector (AEAMFI-Det) for OD in RS images. AEAMFI-Det consists of an adaptive edge aggregation (AEA) module, a feature enhancement module (FEM) embedded in a context-aware cross-attention feature pyramid network (2CA-FPN), and a pyramid squeeze attention (PSA) module. The AEA module employs an edge enhancement mechanism to guide the network to learn spatial multiscale nonlocal dependencies and solve the problem of feature misalignment between the network's focus and the real object. The 2CA-FPN employs level-by-level feature fusion to enhance multiscale feature interactions and effectively mitigate the misalignment between the scales of the extracted features and the scales of real objects. The FEM is designed to capture the local and nonlocal contexts as auxiliary information to enhance the feature representation of information interaction between multiscale features in a cross-attention manner. We introduce the PSA module to establish long-term dependencies between multiscale spaces and channels for better interdependency refinement. Experimental results obtained using the NWPU VHR-10 and DIOR datasets demonstrate the superior performance of AEAMFI-Det in object classification and localization.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001099402800001  
Indexed Date:  
2023-11-28  
Institution Address:  
Zhengzhou Univ Light Ind, Coll Comp & Commun Engn, Zhengzhou 450002, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Jiangsu Collaborat Innovat Ctr Atmospher Environm, Nanjing 210044, Peoples R China  
Institution Address:  
Zhengzhou Univ Light Ind, Engn Training Ctr, Zhengzhou 450002, Peoples R China  
Institution Address:  
Chinese Acad Sci, Chongqing Inst Green & Intelligent Technol, Chongqing 400714, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 105. Multiscale Fusion Network Based on Global Weighting for Hyperspectral

摘要: Feature selection (FS) is an important way to achieve high-precision and efficient classification of hyperspectral remote sensing images. However, most existing FS methods use a fixed scale to extract features and the relationship between spatial and spectral dimensions is ignored. In fact, this correlation is useful for classification. In this article, a multiscale feature fusion network based on global weighting (MSFGW) is proposed in which a global weighting mechanism is explored to catch spatial-spectral information at multiple scales. First, the multiscale feature extraction module composed of group convolution and dilated convolution is utilized to extract the multiscale features. With the increase of the dilation rate, the module takes the spatial differences at varying scales. Second, a 3-D weighting mechanism is used to combine the spatial and spectral correlated information for reducing the interference of homologous and heterologous and boosting the feature discrimination ability. Then, multiscale weighted features are fused to integrate the internal information of all bands at different scales. Finally, the band reconstruction network is used to select representative bands according to their entropy. The experimental results with the state-of-the-art FS algorithms on four widely hyperspectral datasets demonstrate that the features selected by MSFGW have obvious advantages in classification with only a few training samples.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000989479900011  
Indexed Date:  
2023-06-18  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Astronaut, Nanjing 210016, Peoples R China  
Institution:

### 106. YOLOF-F: you only look one-level feature fusion for traffic sign

摘要: This paper proposes a detector that focuses on multi-scale detection problems and effectively enhances the detection performance to solve the problem that is hard to detect minor traffic signs. This detector, called YOLOF-F (you only look one-level feature fusion), is a single-stage detector that extracts multi-scale feature information from a single layer of fusion feature. First, we propose FFM (feature fusion module) to fuse different scales. Next, we offer a new encoder CDE (corner dilated encoder) to enhance the angular point information in the feature map, improve position regression accuracy, and maintain a faster detection speed. Finally, YOLOF-F achieved 74.57% and 77.23% of the AP on the GTSDB and CTSD datasets and reached 32 FPS. Extensive experiments validate that YOLOF-F is faster and more effective than most traffic sign detection methods.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
16  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
16  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:000952181900001  
Indexed Date:  
2023-04-06  
Institution Address:  
Xinjiang Univ, Sch Software, Urumqi 830000, Peoples R China  
Institution Address:  
Key Lab Signal Detect & Proc, Urumqi 83000, Peoples R China  
Institution Address:  
Xinjiang Univ, Key Lab Software Engn, Urumqi 83000, Peoples R China  
Institution:  
Institution:  
Key Lab Signal Detect & Proc  
Institution:

### 107. MFRENet: efficient detection of drone image based on multiscale feature

摘要: The field of object detection in images captured by drones is witnessing a growing surge in research interest. However, because of the abundance of densely packed small objects in the majority of drone images, efficiently detecting dense small objects and achieving accurate classification remain a formidable challenge. To solve the problems mentioned above, we introduce an effective object detection network for drone images based on Multiscale Feature aggregation and Receptive field Expansion (MFRENet). First, we design an effective module named Receptive Field Expanded Feature Extraction Module (RFEFE), which can improve the model's perception ability of objects with irregular shapes and varying sizes. Next, we introduce the Multiscale Cross Stage Parallel Feature Fusion Module (MCSPFF), which integrates the RFEFE module, and then add the Shuffle Attention module to enable MCSPFF to obtain more semantic information. Then, we propose the Extended Simplified Spatial Pyramid Pooling-Fast and Feature Enhancement Module (ESimSPP2FE), which is inspired by the attention mechanism and enhances the features of small objects. Finally, we propose a small target detection head specially used to detect small targets, which enhances the detection ability of our model. Comprehensive experiments are performed on the VisDrone2021-DET dataset, and the proposed model is compared with the baseline YOLOv8m. The experimental results demonstrate that, in comparison to YOLOv8m, the proposed model achieves improvements of 1.9 and 2.7% in mAP and AP50, respectively. The code is available at https://github.com/chenhao-123-sudo/MFRENet-achive.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1433-7541  
eISSN:  
1433-755X  
入藏号:  
WOS:001316837700001  
Indexed Date:  
2024-09-25  
Institution Address:  
Hebei Univ, Sch Cyber Secur & Comp, Baoding 071002, Peoples R China  
Institution Address:  
Hebei Univ, Machine Vis Engn Res Ctr, Baoding 071002, Peoples R China  
Institution:  
Institution:

### 108. Human Pose Estimation Method Based on Optimized Multi-scale Feature

摘要: Human pose estimation is the basis of many tasks in the field of computer vision. Due to the challenge of scale change, the previous human pose estimation network will lose pose information in the process of feature extraction, which makes it difficult to improve the accuracy of human pose estimation. To solve this problem, a parallel network combined with multi-scale feature fusion method is considered to extract features. The human posture estimation method for optimizing feature extraction is divided into two steps: firstly, in the multi-scale feature fusion stage, transpose convolution and mixed dilated convolution are used to reduce the loss of feature information. Secondly, in the feature map output stage, weighted feature maps of different scales are combined to eliminate redundant information, retain posture information, and generate higher quality high-resolution heat map at the same time. Experiments show that the accuracy of this method is improved by 2.1% compared with the advanced method HRnet (High Resolution Net). Experiments show that this method can surpass the existing mainstream human pose estimation methods in accuracy. This method can better meet the challenge of mesoscale change in pedestrian pose estimation, and more accurately locate the key points of small-scale human body in complex scenes.  
摘要:  
人体姿态估计是诸多计算机视觉领域任务的基础,以往的人体姿态估计网络由于尺度变化的挑战会在特征提取的过程中丢失姿态信息,导致人体姿态估计的准确度难以提升。针对该问题考虑以并行网络的方式结合多尺度特征融合方法提取特征。优化特征提取的人体姿态估计方法分为两步:首先在多尺度特征融合阶段利用转置卷积和混合空洞卷积的操作以减少特征信息的丢失,其次在特征图输出阶段有权重的结合不同尺度的特征图来剔除冗余信息、保留姿态信息同时生成更高质量的高分辨率热图。试验表明,此种方法在COCO数据集上的试验结果相比先进方法HRNet(High resolution net)准确率提高了2.1%。通过试验验证本方法在精度方面能够超过现有的主流人体姿态估计方法。该方法能更好地应对行人姿态估计中尺度变化的挑战,更加精确地定位复杂场景中小尺度人体的关键点位置。  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0577-6686  
入藏号:  
CSCD:7855643  
Indexed Date:  
2025-02-09  
Institution Address:  
北京联合大学, 北京信息服务工程重点实验室, 北京 100101, 中国  
Institution Address:  
清华大学车辆与运载学院, 北京 100084, 中国  
Institution Address:  
Beijing Union University, Beijing Key Laboratory of Information Service Engineering, Beijing 100101, China  
Institution Address:  
School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China  
Institution:  
北京联合大学  
Institution:  
清华大学车辆与运载学院  
Institution:  
Beijing Union University  
Institution:  
School of Vehicle and Mobility, Tsinghua University

### 109. Hand-Object Pose Estimation and Reconstruction Based on Signed Distance

摘要: The study of reconstruction of hands and objects from color monocular images has garnered considerable attention in recent years. In existing methods, parametric models are constructed at single scale, and the interaction between hands and objects has not fully be explored. As a result, the multiscale information in 2D images cannot be fully exploited. At the same time, the lack of feature fusion and insufficient utilization of labels also have a great impact on the reconstruction accuracy. To address the limitations, a new framework is proposed, which comprises three key modules. Firstly, a multiscale feature extractor, which generates a multiscale representation of feature, is used to capture the interaction between hand and object more effectively. Secondly, a bridge based on attention has been used to establish the connection between hand and object representations, which facilitates the integration of them. Lastly, a module based on token merge is introduced into the framework, which provides the segmentation representation of object. The experimental results on two datasets, named Obman and DexYCB, demonstrated that the proposed method had good performance and achieved a shape error about 0.121 cm(2) on Obman and 0.40 cm(2) on DexYCB, outperforming the state-of-the-art methods. This study will probably provide the human-computer interaction methods with broader application prospects.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1551-3203  
eISSN:  
1941-0050  
入藏号:  
WOS:001236605600001  
Indexed Date:  
2024-06-06  
Institution Address:  
Fudan Univ, Acad Engn & Technol, Shanghai 200433, Peoples R China  
Institution Address:  
Fudan Univ, Informatizat Off, Shanghai 200433, Peoples R China  
Institution Address:  
Deepearthgo, Shanghai 200000, Peoples R China  
Institution Address:  
Fudan Univ, Acad Engn & Technol, Shanghai 200433, Peoples R China  
Institution Address:  
Shanghai Key Lab Med Image Comp & Comp Assisted In, Shanghai 200032, Peoples R China  
Institution:  
Institution:  
Institution:  
Deepearthgo  
Institution:  
Institution:  
Shanghai Key Lab Med Image Comp & Comp Assisted In

### 110. Few-Shot Remote Sensing Scene Classification via Subspace Based on

摘要: Because of the challenges associated with the difficulty of accurately labeling the remote sensing (RS) scene images and the need to identify new scene classes, few-shot learning has shown significant advantages in addressing the remote sensing scene classification (RSSC) tasks, leading to a growing interest. However, due to the scale variations of targets and irrelevant complex background in scene images, the current few-shot methods exist the following problems: the problem of the extraction capability of feature extractor in the few-shot mechanism; the problem of the separability of few-shot RS scene images classifier. To solve the above problems, an approach, called few-shot RSSC via subspace based on multiscale feature learning is introduced in this work. We first design a multiscale feature learning technique to address scale variations of the targets in the scene images. Concretely, different branches are utilized to learn scene features at various scales. The self-attention mechanism is embedded in each branch to incorporate the understanding of the global information in the different scale features. After that, a multiscale feature fusion operation, incorporating channel attention, will be devised to effectively merge the different scale features, so as to obtain a more precise feature representation of RS scene images. Furthermore, the subspace is utilized to capture the shared characteristics of each category, to reduce the impact of the complex irrelevant backgrounds in the scene images. The results of our experiments conducted on the public available RS scene datasets demonstrate the strong competitiveness of our approach.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001290493100006  
Indexed Date:  
2024-09-09  
Institution Address:  
Chongqing Univ Posts & Telecommun, Sch Commun & Informat Engn, Chongqing 400065, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Sch Intelligent Syst Engn, Shenzhen 518107, Peoples R China  
Institution:  
Institution:

### 111. COVID-19 lesion segmentation based on multi-scale feature fusion and

摘要: A COVID-19 lesion segmentation network based on multi-scale feature fusion and reverse attention (MFFRA) is proposed to overcome the problems of high variability and low contrast between lesion and background in COVID-19 segmentation. The residual network is used as the backbone network to extract features, and the global context aggregation strategy is adopted to integrate different hierarchical features for obtaining rough segmentation results. In addition, the multiscale feature fusion module is added at the bottleneck of the network to enable the ability to segment lesions of different scales using atrous convolutions and multi-kernel pooling. Finally, a novel cascaded reverse attention module is designed to improve the low contrast between normal tissue and lesions based on the detailed features of complementary regions. The proposed method has an accuracy, specificity and sensitivity of 0.714, 0.700, 0.958 on the COVID-19 CT test set, reduces the areas of misdetection and missed detection, and enhances the segmentation ability of fine lesions.  
摘要:  
针对新型冠状病毒肺炎(COVID-19)分割问题中感染区域具有高变异性以及病灶与背景对比度低等问题,提出一种基于多尺度特征融合与反向注意力的COVID-19感染分割网络。首先,利用残差网络作为主干网络进行特征提取,并使用全局上下文聚合策略对不同层次特征进行融合得到粗略的分割结果;其次,在网络瓶颈处添加多尺度特征融合模块,利用空洞卷积与多核池化增强网络分割不同尺度病变的能力;最后,设计一种级联结构的反向注意力模块,利用互补区域的细节特征增强背景与目标的对比度。本文方法在COVID-19 CT分割测试集上的准确率、特异性、灵敏度分别达到0.714、0.700和0.958,误检和漏检区域明显减少,细小病灶的分割能力显著提升。  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1005-202X  
入藏号:  
CSCD:7461907  
Indexed Date:  
2023-08-12  
Institution Address:  
中原工学院电子信息学院;;郑州大学信息工程学院, ;;, 郑州;;郑州, 河南;;河南 450007;;450001, 中国  
Institution Address:  
中原工学院电子信息学院, 郑州, 河南 450007, 中国  
Institution Address:  
河南科技大学第一附属医院儿科, 洛阳, 河南 471000, 中国  
Institution Address:  
School of Electronic and Information, Zhongyuan University of Technology;;School of Information Engineering, Zhengzhou University, ;;, Zhengzhou;;Zhengzhou, ;; 450007;;450001  
Institution Address:  
School of Electronic and Information, Zhongyuan University of Technology, Zhengzhou, Henan 450007, China  
Institution Address:  
Department of Pediatrics, the First Affiliated Hospital of Henan University of Science and Technology, Luoyang, Henan 471000, China  
Institution:  
中原工学院电子信息学院;;郑州大学信息工程学院  
Institution:  
中原工学院电子信息学院  
Institution:  
河南科技大学第一附属医院儿科  
Institution:  
School of Electronic and Information, Zhongyuan University of Technology;;School of Information Engineering, Zhengzhou University  
Institution:  
School of Electronic and Information, Zhongyuan University of Technology  
Institution:  
Department of Pediatrics, the First Affiliated Hospital of Henan University of Science and Technology

### 112. Multi-Scale Feature Fusion Enhancement for Underwater Object Detection

摘要: Underwater object detection (UOD) presents substantial challenges due to the complex visual conditions and the physical properties of light in underwater environments. Small aquatic creatures often congregate in large groups, further complicating the task. To address these challenges, we develop Aqua-DETR, a tailored end-to-end framework for UOD. Our method includes an align-split network to enhance multi-scale feature interaction and fusion for small object identification and a distinction enhancement module using various attention mechanisms to improve ambiguous object identification. Experimental results on four challenging datasets demonstrate that Aqua-DETR outperforms most existing state-of-the-art methods in the UOD task, validating its effectiveness and robustness.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001366537800001  
PubMed ID:  
39598978  
Indexed Date:  
2024-12-07  
Institution Address:  
Guangdong Polytech Normal Univ, Sch Comp Sci, Guangzhou 510665, Peoples R China  
Institution Address:  
Guangdong Polytech Normal Univ, Guangdong Prov Key Lab Intellectual Property & Big, Guangzhou 510665, Peoples R China  
Institution Address:  
Guangdong Polytech Normal Univ, Sch Data Sci & Engn, Heyuan 517583, Peoples R China  
Institution Address:  
Guangdong Univ Finance & Econ, Sch Stat & Math, Guangzhou 510320, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 113. Network and Dataset for Multiscale Remote Sensing Image Change Detection

摘要: Remote sensing image change detection (RSCD) aims to identify differences between remote sensing images of the same location at different times. However, due to the significant variations in the size and appearance of objects in real-world scenes, existing RSCD algorithms often lack strong capabilities in extracting multiscale features, thereby failing to fully capture the characteristics of changes. To address this issue, a multiscale remote sensing change detection network (MSNet) and a multiscale RSCD dataset (MSRS-CD) are proposed. A multiscale convolution module (MSCM) is investigated, and combined with MSCM, an encoder capable of capturing features of different sizes is designed to efficiently extract multiscale semantic change features. A global multiscale feature fusion module is designed to achieve global multiscale feature fusion and obtain multiscale high-level semantic change features. As existing RSCD datasets lack rich scale information and often focus on change targets of specific sizes, a new dataset, MSRS-CD, is constructed. This dataset consists of 842 pairs of images with a resolution of 1024 x 1024 pixels, featuring uniformly distributed change detection target sizes. Comparative experiments are conducted with 10 other state-of-the-art algorithms on the MSRS-CD dataset and another public dataset, LEVIR-CD. Experimental results demonstrate that MSNet achieves the best performance on both datasets, with an F1 score of 75.74% on the MSRS-CD dataset and 91.41% on the LEVIR-CD dataset.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001398107400001  
Indexed Date:  
2025-01-23  
Institution Address:  
Changsha Univ Sci & Technol, Sch Phys & Elect Sci, Changsha 410114, Peoples R China  
Institution:

### 114. Multiscale Feature and Attention Mechanism for Blood Vessel Segmentation

摘要: This study proposes a vascular segmentation network that integrates multiscale feature and a dual attention mechanism to address low segmentation accuracy caused by unsatisfactory segmentation of small retinal vessels and poor vascular connectivity. First, the dilated residual module with introduction of the dual attention mechanism is used to replace the original convolutional layer of U-Net, achieving multiscale extraction of vascular features. Second, a feature fusion module is embedded in the skip connections, reducing information loss during the encoding-decoding process and enhancing vascular connectivity through the adaptive fusion of vascular information. Finally, a hybrid loss function is introduced to assist network training, alleviating the class imbalance problem in retinal vascular images. Experimental results on the DRIVE and CHASE\_DB1 datasets demonstrate that the proposed algorithm achieves an accuracy of 0. 9625 and 0. 9696, respectively. Compared with U-Net, the sensitivity of the proposed algorithm increased by 0.0420 and 0. 0552, and the F1 score increased by 0.0140 and 0. 0342, demonstrating improved segmentation performance.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001438407600009  
Indexed Date:  
2025-03-13  
Institution Address:  
Southwest Univ Sci & Technol, Sch Informat Engn, Mianyang 621000, Sichuan, Peoples R China  
Institution Address:  
Robot Technol Used Special Environm Key Lab Sichua, Mianyang 621000, Sichuan, Peoples R China  
Institution:  
Institution:  
Robot Technol Used Special Environm Key Lab Sichua

### 115. Disparity Estimation Method Based on an Improved ACV Model

摘要: As a hot topic in computer vision, binocular stereo matching has broad applications in various tasks such as distance perception, remote sensing, and autonomous driving. An end- to- end disparity estimation method based on an improved attention concatenation cost volume network is proposed herein to address the challenges of depth discontinuity and inaccurate disparity prediction in boundary regions observed in current methods. First, a multiscale feature fusion network is introduced to combine multiscale feature maps containing rich information from both shallow and deep layers. This approach enhances the fine-grained representation of image details and mitigates the problem of inaccurate disparity prediction in areas with depth discontinuities. Subsequently, a Sobel edge smoothing loss is designed to establish a constraint between the disparity map boundary and the scene's edge contours, alleviating inaccuracies in disparity prediction at the image's target boundaries. Experimental verification of the proposed method on the Sceneflow dataset reveals that the proposed method achieves 0. 467 score in the EPE metric and 1. 51 % in the D1 metric. On the KITTI dataset, the method achieves 1. 44 % score in the 3- All metric and 1. 61 % in the D1- All metric. Compared to the attention concatenation cost volume network, the proposed method shows reduced EPE and D1 scores by 3. 51 % and 5. 63 %, respectively, and reduced 3- All and D1- All metrics by 2. 04 % and 2. 42 %, respectively, demonstrating superior disparity estimation performance.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001397819400006  
Indexed Date:  
2025-01-23  
Institution Address:  
Shanghai Univ Elect Power, Coll Elect & Informat Engn, Shanghai 201306, Peoples R China  
Institution Address:  
Beijing Jiaotong Univ, Sch Elect & Informat Engn, Beijing 100044, Peoples R China  
Institution:  
Institution:

### 116. SANet: A Sea-Land Segmentation Network Via Adaptive Multiscale Feature

摘要: Sea-land segmentation of remote sensing images is of great significance to the dynamic monitoring of coastlines. However, the types of objects in the coastal zone are complex, and their spectra, textures, shapes, and distribution features are different. Therefore, sea-land segmentation for various types of coastlines is still a challenging task. In this article, a scale-adaptive semantic segmentation network, called SANet, is proposed for sea-land segmentation of remote sensing images. SANet has made two innovations on the basis of the classic encoder-decoder structure. First, to integrate the spectral, textural, and semantic features of ground objects at different scales, we designed an adaptive multiscale feature learning module (AML) to replace the conventional serial convolution operation. The AML module mainly contains a multiscale feature extraction unit and an adaptive feature fusion unit. The former can capture the multiscale detailed information and contextual semantic information of objects from an early stage, while the latter can adaptively fuse feature maps of different scales. Second, we adopted the squeeze-and-excitation module to bridge the corresponding layers of the codec so that SANet can selectively emphasize the features of the weak sea-land boundaries. Experiments on a set of Gaofen-1 remote sensing images demonstrated that SANet achieved more accurate segmentation results and obtained sharper boundaries than other methods for various natural and artificial coastlines.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
40  
在中国科学引文数据库中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
5  
被引频次合计:  
44  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000607413900005  
Indexed Date:  
2021-01-27  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao 266590, Peoples R China  
Institution:

### 117. SMN-YOLO: Lightweight YOLOv8-Based Model for Small Object Detection in

摘要: The remote sensing image object detection has advanced significantly; yet, small object detection remains challenging due to their limited size and varying scales. Furthermore, real-world deployment often requires algorithms optimized for fewer parameters and faster inference. To address these issues, we propose SMN-YOLO, a lightweight small object detector based on YOLOv8n. Our approach introduces spatial-channel decoupling downsampling to reduce model size while retaining crucial downsampling information. We also present lightweight and efficient feature pyramid network (LEFPN), a lightweight multiscale feature fusion network incorporating coordinate attention (CA) to capture spatial location cues, enhancing small object detection. In addition, a multiscale feature attention module (MSFAM) further strengthens feature representation. To improve accuracy, we integrate new complete intersection over union (N-CIoU) bounding box regression loss, which minimizes the impact of positional changes on IoU, helping the model focus on low-IoU objects. Experimental results on the vehicle detection in aerial imagery (VEDAI) and AI-based tiny object detection (AI-TOD) datasets show that SMN-YOLO outperforms baseline models with a 3.2% and 2.9% improvement in mean average precision (mAP) at 0.5, respectively, while significantly reducing parameters and only slightly increasing inference time. The proposed model achieves a strong balance between performance and complexity, surpassing several leading detection models.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001442954900006  
Indexed Date:  
2025-03-21  
Institution Address:  
Chinese Acad Sci, Inst Opt & Elect, Chengdu 610209, Peoples R China  
Institution Address:  
Natl Lab Adapt Opt, Chengdu 610209, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 101408, Peoples R China  
Institution:  
Institution:  
Natl Lab Adapt Opt  
Institution:

### 118. Multiscale Feature Fusion for Hyperspectral Marine Oil Spill Image

摘要: Oil spills have always been a threat to the marine ecological environment; thus, it is important to identify and divide oil spill areas on the ocean surface into segments after an oil spill accident occurs to protect the marine ecological environment. However, oil spill area segmentation using ordinary optical images is greatly interfered with by the absorption of light by the deep sea and the distribution of algal organisms on the ocean surface, and it is difficult to improve segmentation accuracy. To address the above problems, a hyperspectral ocean oil spill image segmentation model with multiscale feature fusion (MFFHOSS-Net) is proposed. Specifically, the oil spill segmentation dataset was created using hyperspectral image data from NASA for the Gulf of Mexico oil spill, small-size images after the waveband filtering of the hyperspectral images were generated and the oil spill images were annotated. The model makes full use of having different layers with different characteristics by fusing feature maps of different scales. In addition, an attention mechanism was used to effectively fuse these features to improve the oil spill region segmentation accuracy. A case study, ablation experiments and model evaluation were also carried out in this work. Compared with other models, our proposed method achieved good results according to various evaluation metrics.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2077-1312  
入藏号:  
WOS:001038735700001  
Indexed Date:  
2023-08-12  
Institution Address:  
Chongqing Univ Sci & Technol, Dept Intelligent Technol & Engn, Chongqing 401331, Peoples R China  
Institution Address:  
Chongqing Univ Technol, Dept Comp Sci & Engn, Chongqing 400054, Peoples R China  
Institution:  
Institution:

### 119. Hierarchical Shared Architecture Search for Real-Time Semantic

摘要: Real-time semantic segmentation of remote-sensing images demands a trade-off between speed and accuracy, which makes it challenging. Apart from manually designed networks, researchers seek to adopt neural architecture search (NAS) to discover a real-time semantic segmentation model with optimal performance automatically. Most existing NAS methods stack up no more than two types of searched cells, omitting the characteristics of resolution variation. This article proposes the hierarchical shared architecture search (HAS) method to automatically build a real-time semantic segmentation model for remote sensing images. Our model contains a lightweight backbone and a multiscale feature fusion module. The lightweight backbone is carefully designed with low computational cost. The multiscale feature fusion module is searched using the NAS method, where only the blocks from the same layer share identical cells. Extensive experiments reveal that our searched real-time semantic segmentation model of remote sensing images achieves the state-of-the-art trade-off between accuracy and speed. Specifically, on the LoveDA, Potsdam, and Vaihingen datasets, the searched network achieves 54.5% mIoU, 87.8% mIoU, and 84.1% mIoU, respectively, with an inference speed of 132.7 FPS. Besides, our searched network achieves 72.6% mIoU at 164.0 FPS on the CityScapes dataset and 72.3% mIoU at 186.4 FPS on the CamVid dataset.  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001188382600014  
Indexed Date:  
2024-04-05  
Institution Address:  
Northwestern Polytech Univ, Sch Comp Sci, Shaanxi Prov Key Lab Speech & Image Informat Proc, Xian 710072, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Sch Comp Sci, Natl Engn Lab Integrated Aerosp Ground Ocean Big D, Xian 710072, Peoples R China  
Institution:  
Institution:

### 120. Asymmetric TeacherStudent Feature Pyramid Matching for Industrial

摘要: Visual anomaly detection plays an important role in industrial product quality inspection, but the scarcity of anomaly samples makes the training of supervised models extremely challenging. Many knowledge distillation-based unsupervised detection methods have been proposed recently; however, these anomaly detection methods mostly use similar or identical symmetric structures to build teacher-student (T-S) models, which hinders the different expression of anomalies between T-S networks. To address this issue, this article proposes an asymmetric T-S feature pyramid matching network (ATSN). This network consists of a teacher network, a student network, and a lightweight feature fusion module (FFM). The teacher network is constructed with ResNet18, and the student network comprises a U-shaped network consisting of an encoder and a decoder. In the knowledge transfer process, the encoder acts as an intermediate transmitter of information, and the decoder reconstructs the information transmitted by the encoder into multiscale features. The U-shaped design can effectively improve the accuracy of knowledge transfer, thus increasing the differences in anomalous representation via T-S feature pyramid matching. Additionally, an FFM is used to fuse the correct information from the multiscale feature representation in the encoder and the decoder, thereby compensating for the loss of key information in the decoder and further increasing the difference in anomalous representation. Experiments on the MVTec AD dataset show that our proposed method surpasses the current state-of-the-art methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001132683400205  
Indexed Date:  
2024-02-02  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Software, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Software, Jiangsu Prov Engn Res Ctr Adv Comp & Intelligent, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Jiangsu Collaborat Innovat Ctr Atmospher Environm, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ, State Key Lab Novel Software Technol, Nanjing 210023, Peoples R China  
Institution Address:  
Nanjing Normal Univ Special Educ, Coll Math & Informat Sci, Nanjing 210038, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 121. Multiscale feature fusion network for monocular complex hand pose

摘要: Hand pose estimation based on a single RGB image has low accuracy due to the complexity of the pose, local self-similarity of finger features, and occlusion. A multiscale feature fusion network (MS-FF) for monocular vision gesture pose estimation is proposed to address this problem. The network can take full advantage of different channel information to enhance important gesture information, and it can simultaneously extract features from feature maps of different resolutions to obtain as much detailed feature information and deep semantic information as possible. The feature maps are merged to obtain the hand pose results. The InterHand2.6M dataset and Rendered Handpose Dataset (RHD) are used to train the MS-FF. Compared with the other methods, the MS-FF obtains the smallest average error of hand joints, verifying its effectiveness.  
The authors proposed an MS-FF for monocular visual hand pose estimation. To effectively process the detailed information of occluded edges and fingertips, the network can extract information of different levels from feature maps of different resolutions to more accurately estimate hand poses. A channel conversion module adjusts the weights of channels. To make full use of both the edge detail characteristics of the images and deep semantic information, a global regression module fuses feature maps of different resolutions. An optimization procedure corrects some joints that are not returned to the correct position. Higher accuracy and robustness were achieved using the proposed method. Experiments verified the effectiveness of the MS-FF.image  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0013-5194  
eISSN:  
1350-911X  
入藏号:  
WOS:001129218300001  
Indexed Date:  
2024-02-16  
Institution Address:  
Guangdong Engn Polytech, Zhan Zhi, Peoples R China  
Institution Address:  
South China Normal Univ, Luo Guang, Peoples R China  
Institution:  
Guangdong Engn Polytech  
Institution:

### 122. Adaptive multi-scale feature fusion based U-net for fracture

摘要: Accurate segmentation of fractures in coal rock CT images is important for the development of coalbed methane. However, due to the large variation of fracture scale and the similarity of gray values between weak fractures and the surrounding matrix, it remains a challenging task. And there is no published dataset of coal rock, which make the task even harder. In this paper, a novel adaptive multi-scale feature fusion method based on U-net (AMSFF-U-net) is proposed for fracture segmentation in coal rock CT images. Specifically, encoder and decoder path consist of residual blocks (ReBlock), respectively. The attention skip concatenation (ASC) module is proposed to capture more representative and distinguishing features by combining the high-level and low-level features of adjacent layers. The adaptive multi-scale feature fusion (AMSFF) module is presented to adaptively fuse different scale feature maps of encoder path; it can effectively capture rich multi-scale features. In response to the lack of coal rock fractures training data, we applied a set of comprehensive data augmentation operations to increase the diversity of training samples. These extensive experiments are conducted via seven state-of-the-art methods (i.e., FCEM, U-net, Res-Unet, Unet++, MSN-Net, WRAU-Net and ours). The experiment results demonstrate that the proposed AMSFF-U-net can achieve better segmentation performance in our works, particularly for weak fractures and tiny scale fractures.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
被引频次合计:  
7  
ISSN:  
1064-1246  
eISSN:  
1875-8967  
入藏号:  
WOS:000771196600061  
Indexed Date:  
2022-04-01  
Institution Address:  
China Univ Min & Technol, Sch Mech Elect & Informat Engn, Ding 11 Xueyuan Rd, Beijing 100083, Peoples R China  
Institution:

### 123. Efficient feature fusion network based on center and scale prediction

摘要: Center and scale prediction (CSP) is an anchor-free pedestrian detector with good performance. However, there are lots of parameters in the detector, which seriously limits the speed. In this paper, a new network is designed for the improvement of the detector speed, which contains less parameters, named Feature Fusion: Center and Scale Prediction (F-CSP). F-CSP fuses multi-scale feature maps with two efficient feature fusion networks: Feature Pyramid Networks (FPN) and Balanced Feature Pyramid (BFP). Specifically, FPN is used to reduce the channel of feature maps, and BFP is used to fuse multiple feature maps into a single one. This way, the proposed detector achieves competitive accuracy and higher speed on the challenging pedestrian detection benchmark. The performance of F-CSP is demonstrated on the Caltech dataset. Compared with CSP, under the premise of ensuring accuracy, the speed is increased from 45.1 to 32.9 ms/img.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:000828445600002  
Indexed Date:  
2022-07-28  
Institution Address:  
Tianjin Univ, Sch Elect & Informat Engn, Tianjin 300072, Peoples R China  
Institution:

### 124. CAA-YOLO: Combined-Attention-Augmented YOLO for Infrared Ocean Ships

摘要: Infrared ocean ships detection still faces great challenges due to the low signal-to-noise ratio and low spatial resolution resulting in a severe lack of texture details for small infrared targets, as well as the distribution of the extremely multiscale ships. In this paper, we propose a CAA-YOLO to alleviate the problems. In this study, to highlight and preserve features of small targets, we apply a high-resolution feature layer (P2) to better use shallow details and the location information. In order to suppress the shallow noise of the P2 layer and further enhance the feature extraction capability, we introduce a TA module into the backbone. Moreover, we design a new feature fusion method to capture the long-range contextual information of small targets and propose a combined attention mechanism to enhance the ability of the feature fusion while suppressing the noise interference caused by the shallow feature layers. We conduct a detailed study of the algorithm based on a marine infrared dataset to verify the effectiveness of our algorithm, in which the AP and AR of small targets increase by 5.63% and 9.01%, respectively, and the mAP increases by 3.4% compared to that of YOLOv5.  
在中国科学引文数据库中的被引频次:  
6  
在 BIOSIS Citation Index 中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
29  
被引频次合计:  
35  
eISSN:  
1424-8220  
入藏号:  
WOS:000801950200001  
PubMed ID:  
35632198  
Indexed Date:  
2022-06-07  
Institution Address:  
Beijing Inst Technol, Sch Life Sci, Beijing 100081, Peoples R China  
Institution Address:  
Minist Ind & Informat Technol, Key Lab Convergence Med Engn Syst & Healthcare Te, Beijing 100081, Peoples R China  
Institution Address:  
Beijing Inst Technol, Sch Med Technoloy, Beijing 100081, Peoples R China  
Institution:  
Institution:  
Minist Ind & Informat Technol  
Institution:

### 125. Optimally-Weighted Multi-Scale Local Feature Fusion Network for Driver

摘要: Distracted driving is one of the main contributors to traffic accidents. In this work, we propose a novel multi-scale local feature fusion network for image-based distracted driver detection. Since the driver is the most important part to infer the distracted driver actions in a single image, our proposed method first detects the driver's body using person detection. Then capture abundant local body features after a repeated multi-scale feature fusion module. In addition to the features extracted from the whole image, our network also include the important feature of local body feature. The global feature and local feature are finally fused by an OAWS(optimally-weighted strategy). The experimental result shows that our methods achieve comparative performance on our own HY Large Vehicle Driver Dataset and the public AUC Driver Distracted Dataset.  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:000895896100001  
Indexed Date:  
2022-12-23  
Institution Address:  
Huaiyin Inst Technol, Coll Comp & Software Engn, Lab Internet Things & Mobile Internet Technol Jian, Huaian 223001, Peoples R China  
Institution:

### 126. Steel strip surface defect detection based on multiscale feature sensing

摘要: A surface defect detection method for hot-rolled steel strips was proposed to address the challenges of detecting small target defects, significant differences in morphology, and unclear defect characteristics. This method is based on multiscale feature perception and adaptive feature fusion. First, based on the spatial distribution characteristics of the steel strip image, redundant background interference is removed using automatic gamma correction and Otsu thresholding. Second, based on the characteristics of surface defects in steel strips, this paper proposes TDB-YOLO (YOLO with a small target detection layer), a Bidirectional Feature Pyramid Network (BiFPN), and Double Cross Stage Partial (CSP) Bottleneck with three convolutions (DC3). To detect small object defects, a small target detection layer with a smaller receptive field focuses on fine-grained features, reducing the model's probability of missed detection. In terms of feature extraction, DC3 enhances the interaction of feature information from different spatial scales, enabling the model to effectively handle features of varying scales. In terms of feature fusion, the BiFPN is used to adaptively fuse deep-level and shallow-level feature information, enhancing the semantic richness of the feature information. Ultimately, the proposed model in this paper achieved an accuracy of 90.3% and a recall rate of 88.0% for surface defects in steel strips. The mean average precision was 90.4%, and the frames per second was 33. The detection performance of this model outperformed those of other detection models, demonstrating its ability to effectively meet the real-time detection requirements of surface defects in industrial scenarios on steel strips.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
eISSN:  
2158-3226  
入藏号:  
WOS:001195792200001  
Indexed Date:  
2024-04-07  
Institution Address:  
Chongqing Univ Technol, Coll Mech Engn, Chongqing, Peoples R China  
Institution:

### 127. RMFENet: Refined Multiscale Feature Enhancement Network for

摘要: Automatic underwater vehicle (AUV) object detection in sonar images has drawn widespread interest for its practical applications. Nevertheless, accurate detection remains challenging due to the scale imbalance and arbitrary orientations of underwater objects with extreme aspect ratios. In this article, we introduce a refined multiscale feature enhancement network (RMFENet) framework as a novel approach to achieve high-precision performance in detecting arbitrary-oriented underwater objects. First, we develop a data preprocessing module to adaptively adjust large-size sonar images to the appropriate size and direction angle. Second, we construct a composite backbone network (CB-Net) that refines the integration of high-level and low-level features in sonar images and enhances the correlation of different features. Third, the shuffle convolutional block attention mechanism (Shuffle-CBAM) mechanism and the diverse multiscale feature fusion (DMSFF) block further enhance the model feature extraction ability. Finally, to generate more accurate rotation region proposals and anchors, we propose a rotated intersection over union (RIoU) to solve inaccurate object region localization problems and high computational complexity in existing methods. Besides, our model is constructed on two sonar image datasets for oriented underwater object detection with different spatial resolutions, diverse object categories, and bounding box annotation. Extensive experiments demonstrate that our method can achieve state-of-the-art performance on the constructed sonar image object detection datasets while keeping better robustness.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001161161900085  
Indexed Date:  
2024-03-29  
Institution Address:  
Air Force Engn Univ, Sch Informat & Nav, Xian 710082, Peoples R China  
Institution Address:  
Xijing Univ, Sch Elect Informat, Xian 710123, Peoples R China  
Institution:  
Institution:

### 128. A Variable-scale VS-UNet Model for Road Crack Detection

摘要: Existing image segmentation algorithms face challenges related to low detection accuracy and a lack of specificity in crack detection. To address these challenges, this paper proposes an extended LG-Block module Extend-LG Block, which leverages a multi-scale feature fusion method. This module consists of multiple parallel dilated convolutions with different expansion rates. The number of branches and the expansion rate of dilated convolutions can be adjusted by parameters to change the size of its receptive field, and then extract and fuse crack features of different scales. By comparing the advantages and disadvantages of the network using a multi-scale feature fusion module in the deep layer and the network using a fixed scale structure for multi-scale feature fusion, a U-Net model with a variable scale structure named VS-UNet is proposed. The basic convolution Block in the UNet network is replaced by multiple Extend-LG blocks with different parameters. This structure performs multi-scale feature fusion in the shallow layer of the network, and the scale extracted by the multi-scale feature fusion module gradually decreases with the deepening of the network layer. This structure not only strengthens the detail feature extraction ability of the image while maintaining the original abstract feature extraction ability but also avoids the problem of increasing network parameters caused by the increase of convolution. Experiments are carried out on the DeepCrack dataset and CFD dataset. The results show that compared with the other two structures and methods, the proposed network with variable scale structure has higher detection accuracy and better segmentation effect for cracks of various sizes in visual experimental comparison. Finally, compared with other image segmentation algorithms, all indicators are improved to a certain extent compared with UNet, which proves the effectiveness of the improved network.  
摘要:  
为解决目前现有的图像分割算法存在检测精度低、对裂缝检测缺乏针对性等问题,采用多尺度特征融合方法,提出一种扩展LG Block模块Extend-LG Block,其由多个并行不同膨胀率的空洞卷积组成.通过参数可调节分支数量和空洞卷积膨胀率,从而改变其感受野大小,进而提取和融合不同尺度的裂缝特征.对比在深层使用多尺度特征融合模块的网络以及使用固定尺度结构进行多尺度特征融合的网络的优劣,提出一种变尺度结构的UNet模型VS-UNet,使用多个不同参数的Extend-LG Block替换UNet网络中的基本卷积块.该结构在网络浅层进行多尺度特征融合,多尺度特征融合模块提取的尺度随网络层加深逐渐减少.此结构在加强图像的细节特征提取能力的同时保持原有的抽象特征提取能力,还可避免网络参数的增加.在DeepCrack数据集以及CFD数据集上进行实验验证,结果表明,相较于其他两种结构和方法,提出的变尺度结构的网络在有更高检测精度的同时,在可视化实验对比上对各种大小的裂缝有更好的分割效果.最后与其他图像分割算法进行对比,各项指标与UNet相比均有一定程度提升,证明了网络改进的有效性.研究结果可为进一步提升道路裂缝检测效果提供参考.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1674-2974  
入藏号:  
CSCD:7755178  
Indexed Date:  
2024-08-30  
Institution Address:  
石家庄铁道大学;;石家庄铁道大学信息科学与技术学院, 河北省省部共建交通工程结构力学行为与系统安全国家重点实验室;;, 石家庄;;石家庄, 河北;;河北 050043;;050043, 中国  
Institution Address:  
石家庄铁道大学信息科学与技术学院, 石家庄, 河北 050043, 中国  
Institution Address:  
Shijiazhuang Tiedao University;;School of Information Science and Technology, Shijiazhuang Tiedao University, State Key Laboratory of Mechanical Behavior and System Safety of Traffic Engineering Structures;;, Shijiazhuang;;Shijiazhuang, ;; 050043;;050043  
Institution Address:  
School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, Hebei 050043, China  
Institution:  
石家庄铁道大学;;石家庄铁道大学信息科学与技术学院  
Institution:  
石家庄铁道大学信息科学与技术学院  
Institution:  
Shijiazhuang Tiedao University;;School of Information Science and Technology, Shijiazhuang Tiedao University  
Institution:  
School of Information Science and Technology, Shijiazhuang Tiedao University

### 129. Maritime Ship Detection Method for Satellite Images Based on Multiscale

摘要: Remote sensing ship recognition technology is one of the important research areas for achieving ocean security monitoring. However, maritime ship targets features such as multiscale, arbitrary direction, and dense arrangement, and their imaging is often influenced by factors such as sea fog and sea surface background. It is challenging for fast and accurate detection of remote sensing ships. We propose a new ship detection algorithm YOLO-remote sensing ship detection (YOLO-RSSD) based on YOLOv5, which effectively improves the accuracy of ship detection while ensuring detection speed. It is attributed to our proposed four optimization measures in this model. First, the K-means++ algorithm is introduced in the data preprocessing part to improve the clustering effect and obtain candidate box sizes suitable for multiscale ship dataset. Then, an improved bidirectional feature pyramid network structure is embedded in the feature fusion part to achieve cross-layer multiscale weighted feature fusion. Third, the original bounding box regression loss function is replaced with the EIoU loss, which is effective for accelerating the convergence and improving the regression accuracy of predicted boxes. Finally, a channel attention mechanism is introduced in the convolutional unit to enhance the model's ability to capture ship features. Experimental results show that the YOLO-RSSD model achieves a detection accuracy of 96.1% for remote sensing ships, which is 4.3% higher than the original YOLOv5 network. In addition, YOLO-RSSD performs good robustness and generalization ability. This means that our method has high practical value and provides a new solution for the analysis and application of remote sensing images.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
20  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
21  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001037835600009  
Indexed Date:  
2023-08-12  
Institution Address:  
Tianjin Univ Technol, Engn Res Ctr Learning Based Intelligent Syst, Key Lab Comp Vis & Syst, Minist Educ, Tianjin 300384, Peoples R China  
Institution Address:  
Tianjin Univ Technol, Sch Comp Sci & Engn, Tianjin 300384, Peoples R China  
Institution Address:  
Tiangong Univ, Sch Elect & Informat Engn, Tianjin Key Lab Optoelect Detect Technol & Syst, Tianjin 300387, Peoples R China  
Institution Address:  
Nankai Univ, Inst Modern Opt, Tianjin Key Lab ofMicro scale Opt Informat Sci & T, Tianjin 300350, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 130. Image Dehazing via Double-layer Vision and Multi-scale Attention Fusion

摘要: Aiming at the problem that the existing dehazing algorithms lack attention to the noise concentration in different regions of the hazy image and the distinction between far and near features, this paper proposes a new generative adversarial network model. In the model, two UNet3 + networks are used to realize the full-scale jump connection and depth supervision, and multi-scale feature fusion is used to extract the high and low-level semantics in different scale feature images. The addition of deep supervision can better learn the near-far level representation in the image. At the same time, the multi-scale pyramid feature fusion module integrating the self-attention mechanism is added to the generator structure to better retain the multi-scale structure information of the feature map and improve the attention to different haze concentration regions. The experimental results show that the algorithm network can obtain better visual effects than other advanced algorithms such as BPPNET on NTIRE 2020, NTIRE 2021, O-Haze datasets, and Dense-Haze datasets. The peak signal-to-noise ratio and structural similarity index on the Dense-Haze dataset are, respectively, 24.82 and 0.769.  
摘要:  
针对现有去雾算法缺乏对雾霾图像不同区域噪音浓度的关注以及远近景特征的区分问题,本文提出了一种新的生成对抗网络模型.模型中通过两个UNet3+网络实现全尺度的跳跃连接和深度监督,使用多尺度融合的方法结合不同尺度特征图中的高低级语义;而深度监督的加入可以更好地学习图像中的远近层次表示.同时在生成器结构中加入融合改进自注意力机制的多尺度金字塔特征融合模块,以便更好地保留特征图的多尺度结构信息,并且提高了对不同雾霾浓度区域的关注度.实验结果显示,在NTIRE 2020、NTIRE 2021、O-Haze数据集和Dense-Haze数据集上,本文所提出的算法网络相比BPPNET等其他先进算法可以得到更好的视觉效果,在Dense-Haze数据集上,峰值信噪比和结构相似性指数分别达到24.82和0.769.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1674-2974  
入藏号:  
CSCD:7421882  
Indexed Date:  
2023-06-16  
Institution Address:  
兰州交通大学电子与信息工程学院, 兰州, 甘肃 730070, 中国  
Institution Address:  
College of Electronics and Information Engineering,Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China  
Institution:  
兰州交通大学电子与信息工程学院  
Institution:  
College of Electronics and Information Engineering,Lanzhou Jiaotong University

### 131. Multi-Scale Object Detection with the Pixel Attention Mechanism in a

摘要: The object detection task is usually affected by complex backgrounds. In this paper, a new image object detection method is proposed, which can perform multi-feature selection on multi-scale feature maps. By this method, a bidirectional multi-scale feature fusion network was designed to fuse semantic features and shallow features to improve the detection effects of small objects in complex backgrounds. When the shallow features are transferred to the top layer, a bottom-up path is added to reduce the number of network layers experienced by the feature fusion network, reducing the loss of shallow features. In addition, a multi-feature selection module based on the attention mechanism is used to minimize the interference of useless information in subsequent classification and regression, allowing the network to adaptively focus on appropriate information for classification or regression to improve detection accuracy. Because the traditional five-parameter regression method has severe boundary problems when predicting objects with large aspect ratios, the proposed network treats angle prediction as a classification task. The experimental results on the DOTA dataset, the self-made DOTA-GF dataset and the HRSC 2016 dataset show that, compared with several popular object detection algorithms, the proposed method has certain advantages in detection accuracy.  
Web of Science 核心合集中的 "被引频次":  
8  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
eISSN:  
2072-4292  
入藏号:  
WOS:000845303900001  
Indexed Date:  
2022-09-03  
Institution Address:  
Wuhan Univ, Sch Elect Informat, Wuhan 430064, Peoples R China  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & R, Wuhan 430079, Peoples R China  
Institution Address:  
China Aerosp Sci & Technol Corp, Aerosp Syst Dev Res Ctr, Beijing 100094, Peoples R China  
Institution Address:  
Qian Xuesen Lab Space Technol, Beijing 100094, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Qian Xuesen Lab Space Technol

### 132. ECG Signals Deep Compressive Sensing Framework Based on Multiscale

摘要: Electrocardiogram (ECG) is nowadays an important technology to be applied in the clinical diagnosis for the detection of the heart disease. But the large storage and high-burden transmission of the ECG data is a challenge. Therefore, the compressive sensing (CS) is appropriate to deal with those signals for it can compress and sample the signal at the same time. In order to get rid of the constraints in the traditional CS methods, we propose a compressive sensing framework based on multiscale feature fusion and SE block. In the compression process we use sequential convolutional layers instead of the traditional compressive sensing using measurement matrix projection for ECG signals. In the reconstruction process, the multi-scale feature fusion method is first used to fuse multiple feature maps output from the convolution layer to better extract signal features. Subsequently, Squeeze-and-Excitation (SE) block is used to further enhance the feature representation. Finally, sequence modeling of the ECG signal is performed using LSTM to obtain the reconstructed signal. The results show that the proposed method performs well on various datasets and evaluation metrics, in the case of SR = 0.4, the PRD and SNR of the experiments on the MIT-BIH Arrhythmia database are 1.55% and 37.66dB, respectively. The PRD and SNR of the experiments on the Non-Invasive Fetal ECG Arrhythmia Database were 2.48% and 34.57dB, respectively, which were the lowest among all the comparison methods, indicating that the proposed method has good ECG signal processing capability.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
2169-3536  
入藏号:  
WOS:001081635300001  
Indexed Date:  
2023-10-26  
Institution Address:  
Jiangxi Agr Univ, Sch Software, Nanchang 330045, Peoples R China  
Institution Address:  
Jiangxi Agr Univ, Sch Comp & Informat Engn, Nanchang 330045, Peoples R China  
Institution:  
Institution:

### 133. Ship fine-grained classification network based on multi-scale feature

摘要: Ships are essential maritime transportation carriers and military targets. Utilizing remote sensing technology to accurately and automatically identify and classify them has broad application prospects and significant practical implications. The challenge in fine-grained ship image classification lies in the subtle differences between various types of ships and significant variations within the same type. This challenge is further compounded by the scarcity of available ship datasets. In order to tackle these issues, we have put forward a ship image classification model that utilizes multi-scale feature fusion for more precise analysis. This research used a multi-scale feature enhancement module to extract feature representations from multi-scale ship images and enhance the intra-class correlation in ships. Subsequently, the multi-scale feature fusion module combined the features to get global correlations of ship images as well as local features. Additionally, this research established a ship dataset comprising six common categories. Ultimately, we assessed the suggested approach on both the FGSCR-42 dataset and our own ship dataset. The results indicate that when compared to ResNet-50, this method significantly improved the fine-grained classification accuracy of ships. The accuracy, precision, recall rate, and the F1Score index reached 98.96%, 97.58%, 97.54%, and 97.56%, respectively.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0029-8018  
eISSN:  
1873-5258  
入藏号:  
WOS:001393349900001  
Indexed Date:  
2025-01-15  
Institution Address:  
Shanghai Maritime Univ, Coll Ocean Sci & Engn, Shanghai, Peoples R China  
Institution Address:  
Shanghai Maritime Univ, Coll Transport & Commun, Shanghai, Peoples R China  
Institution Address:  
Shanghai Maritime Univ, Merchant Marine Coll, Shanghai, Peoples R China  
Institution:  
Institution:  
Institution:

### 134. Optimization Strategy for X-Ray Generation and Countermeasure Fusion of

摘要: During the non-contact flaw detection of a rust-covered bronze mirror, X-ray imaging typically fails to reveal the extent of damage due to the thickness difference between the mirror edge and core. In this study, the X-ray signal from a bronze mirror was used as an input to construct a generative confrontation fusion network. An optimization strategy that enhances the bronze mirror X-ray information fusion was designed to address the reconstruction blur caused by the L-2 loss and gradient operator, and the expression of multiscale feature details, such as textures and cracks. By utilizing the feature learning process of the L-2,L-1/2 loss regularization generator, the smoothing of the data that was generated using the L-2 loss was improved; moreover, the Laplacian L-tex pattern loss was defined to strengthen the effect of training network on the extraction of decorations and diseases. Furthermore, a multiscale feature fusion module was added to the training network to improve the quality of the generated information. Thus, considering the experimental comparison involving seven fusion methods, the cross entropy value of the proposed algorithm in two of the five groups is poor. However, the values are optimal in the control data, including entropy, average gradient, spatial frequency, joint entropy, and non-reference feature mutual information. This can effectively reveal the detection information of the bronze mirror during X-ray flaw detection.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001154750000051  
Indexed Date:  
2024-03-13  
Institution Address:  
Xian Univ Architecture & Technol, Coll Informat & Control Engn, Xian 710055, Shaanxi, Peoples R China  
Institution Address:  
Shaanxi Inst Preservat Cultural Heritage, Xian 710075, Shaanxi, Peoples R China  
Institution:  
Institution:  
Shaanxi Inst Preservat Cultural Heritage

### 135. MFFE: Multi-scale Feature Fusion Enhanced Net for image dehazing

摘要: It is very challenging to perform dehazing operations on single haze images taken due to information degradation, for which we propose a novel single-image dehazing network based on U-Net. Most existing image dehazing algorithms only focus on whether they can remove the haze, but ignore the quality of the final dehazed image, which can lead to problems such as loss of image detail information and texture blurring, affecting the final recovery. We use the enhancement and error feedback mechanisms inside the super-resolution algorithm to gradually recover the haze-free images by introducing the Strengthen-Operate subtract strategy in the decoder and demonstrate their effectiveness for the dehazing problem. Meanwhile, to solve the problem of missing information and utilization in the traditional U-Net structure, we design a multiscale feature fusion module that can effectively compensate for the missing contextual information and make full use of the disjoint features. In addition, an improved local binary pattern and SE attention mechanism are used to help the network obtain clearer details and texture images, and the accuracy of the enhanced dehazing network is improved using residual learning. We analyze the effectiveness of the proposed algorithm and show through extensive evaluation that our proposed model can be effectively used for single image dehazing and also has good performance compared to state-of-the-art methods on the test dataset.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0923-5965  
eISSN:  
1879-2677  
入藏号:  
WOS:000805032000001  
Indexed Date:  
2022-06-11  
Institution Address:  
Shandong Technol & Business Univ, Sch Informat & Elect Engn, Yantai 264005, Peoples R China  
Institution Address:  
Shandong Technol & Business Univ, Sch Comp Sci & Technol, Yantai 264005, Peoples R China  
Institution Address:  
Coinnovat Ctr Shandong Coll & Univ, Future Intelligent Comp, Yantai 264005, Peoples R China  
Institution:  
Institution:  
Institution:  
Coinnovat Ctr Shandong Coll & Univ

### 136. Lightweight Apple-Leaf Pathological Recognition Based on Multiscale

摘要: The occurrence of apple leaf diseases has a significant impact on apple quality and yield. Disease monitoring is therefore an important measure to ensure the healthy development of the apple industry. Based on the ResNet structure, a lightweight disease recognition model based on multiscale feature fusion is proposed. First, the feature fusion mechanism is used to extract and fuse the high-dimensional and low-dimensional features of the network, strengthen the transmission of semantic information between convolution layers, and enhance the ability to distinguish subtle lesions. Next, multiscale depth separable convolution is added to extract disease features of different scales by using multi-scale convolution kernel structure, which improves the richness of features and restricts the parameters of the model. Finally, a dataset containing five kinds of apple leaf diseases is used to verify the effectiveness of the proposed method. The experimental results show that the recognition accuracy of the model is 98. 05%, and that the number and calculation of the model network are only 4. 02 MB and 0. 92 GB, respectively. Compared with other models, it also has advantages, and can provide a new scheme for the accurate identification of diseases and pests in agricultural automation.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1006-4125  
入藏号:  
WOS:001154750000009  
Indexed Date:  
2024-03-13  
Institution Address:  
Guizhou Univ, Coll Big Data & Informat Engn, Guiyang 550025, Guizhou, Peoples R China  
Institution Address:  
Minist Educ, Semicond Power Device Reliabil Engn Res Ctr, Guiyang 550025, Guizhou, Peoples R China  
Institution Address:  
Guizhou Univ, Sch Mech Engn, Guiyang 550025, Guizhou, Peoples R China  
Institution Address:  
Guizhou Univ, Sch Comp Sci & Technol, Guiyang 550025, Guizhou, Peoples R China  
Institution:  
Institution:  
Minist Educ  
Institution:  
Institution:

### 137. MFFAMM: A Small Object Detection with Multi-Scale Feature Fusion and

摘要: Aiming at the low detection accuracy and poor positioning for small objects of single-stage object detection algorithms, we improve the backbone network of SSD (Single Shot MultiBox Detector) and present an improved SSD model based on multi-scale feature fusion and attention mechanism module in this paper. Firstly, we enhance the feature extraction ability of the shallow network through the feature fusion method that is beneficial to small object recognition. Secondly, the RFB (Receptive Field block) is used to expand the object's receptive field and extract richer semantic information. After feature fusion, the attention mechanism module is added to enhance the feature information of important objects and suppress irrelevant other information. The experimental results show that our algorithm achieves 80.7% and 51.8% mAP on the PASCAL VOC 2007 classic dataset and MS COCO 2017 dataset, which are 3.2% and 10.6% higher than the original SSD algorithm. Our algorithm greatly improves the accuracy of object detection and meets the requirements of real-time.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
9  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
eISSN:  
2076-3417  
入藏号:  
WOS:000858026400001  
Indexed Date:  
2022-09-29  
Institution Address:  
Chongqing Univ Posts & Telecommun, Coll Comp Sci & Technol, Chongqing 400065, Peoples R China  
Institution Address:  
Southwest Comp Co Ltd, Inst Informat Technol, Chongqing 400065, Peoples R China  
Institution:  
Institution:  
Southwest Comp Co Ltd

### 138. GDE-Pose: A Real-Time Adaptive Compression and Multi-Scale Dynamic

摘要: This paper introduces a novel lightweight pose estimation model, GDE-pose, which addresses the current trade-off between accuracy and computational efficiency in existing models. GDE-pose builds upon the baseline YOLO-pose model by incorporating Ghost Bottleneck, a Dynamic Feature Fusion Module (DFFM), and ECA Attention to achieve more effective feature representation and selection. The Ghost Bottleneck reduces computational complexity, DFFM enhances multi-scale feature fusion, and ECA Attention optimizes the selection of key features. GDE-pose improves pose estimation accuracy while preserving real-time performance. Experimental results demonstrate that GDE-pose achieves higher accuracy on the COCO dataset, with a substantial reduction in parameters, over 80% fewer FLOPs, and an increased inference speed of 31 FPS, underscoring its exceptional lightweight and real-time capabilities. Ablation studies confirm the independent contribution of each module to the model's overall performance. GDE-pose's design highlights its broad applicability in real-time pose estimation tasks.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001377748800001  
Indexed Date:  
2024-12-25  
Institution Address:  
City Univ Macau, Fac Data Sci, Taipa 999078, Macau, Peoples R China  
Institution:

### 139. Gating attention convolutional networks with dense connection for

摘要: Automatic detection of pavement cracks is an important task to ensure pavement safety. However, noise and uneven cracks in real pavement images present great challenges for crack detection. To address these issues, we propose a novel pavement crack detection network model with a densely connected architecture and a gating attention mechanism. Based on SegNet, the proposed network utilizes an atrous convolutional dense connection module (AD-block) to efficiently extract crack features with different structures in the encoding stage. In addition, for locating crack pixels in the decoding stage, a new gating attention unit (GAU) is designed that can suppress the background noise and accurately locate the crack pixels. Finally, by means of a new multiscale feature fusion (MFF) module, the side outputs are aggregated to obtain the final prediction results. Evaluations of the public DeepCrack, CFD, and Crack500 datasets show that our method achieves better performance than other recent approaches.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
0942-4962  
eISSN:  
1432-1882  
入藏号:  
WOS:000868374900001  
Indexed Date:  
2022-10-26  
Institution Address:  
Chongqing Univ Posts & Telecommun, Chongqing, Peoples R China  
Institution:

### 140. Multi-Scale Feature Fusion of Covariance Pooling Networks for

摘要: Multi-scale feature fusion techniques and covariance pooling have been shown to have positive implications for completing computer vision tasks, including fine-grained image classification. However, existing algorithms that use multi-scale feature fusion techniques for fine-grained classification tend to consider only the first-order information of the features, failing to capture more discriminative features. Likewise, existing fine-grained classification algorithms using covariance pooling tend to focus only on the correlation between feature channels without considering how to better capture the global and local features of the image. Therefore, this paper proposes a multi-scale covariance pooling network (MSCPN) that can capture and better fuse features at different scales to generate more representative features. Experimental results on the CUB200 and MIT indoor67 datasets achieve state-of-the-art performance (CUB200: 94.31% and MIT indoor67: 92.11%).  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
eISSN:  
1424-8220  
入藏号:  
WOS:000977692800001  
PubMed ID:  
37112311  
Indexed Date:  
2023-05-11  
Institution Address:  
Soochow Univ, Sch Rail Transportat, Suzhou 215131, Peoples R China  
Institution Address:  
ByteDance Inc, Moutainview, CA 94042 USA  
Institution:  
Institution:  
ByteDance Inc

### 141. Development of a lightweight cross-scale decoupling feature fusion

摘要: Permanent magnet motors may suffer from imperceptible localized demagnetization due to surface damage during operation. This paper proposes a lightweight cross-scale decoupling feature fusion network (LCDFFN) for permanent magnets defect detection, which introduces separable convolution in the backbone and incorporates the adaptive multi-scale feature fusion with interactive attention mechanism to enhance small target detection. The adaptive attention module (AAM) and feature enhancement module based on dilated convolution are introduced in the feature fusion network, improving multi-scale object detection. Finally, the decoupled feature prediction network outputs the defect identification feature map. In addition, we also propose a novel fuzzy intersection over the union loss function. LCDFFN achieves a mAP@0.5 of 97.8%, with 108.4M parameters and a detection speed of 113.62 FPS. The proposed method significantly improves defect image detection for permanent magnets and is highly practical for industrial production.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001423406400001  
Indexed Date:  
2025-02-23  
Institution Address:  
Hefei Univ Technol, Sch Instrument Sci & Optoelect Engn, Anhui Prov Key Lab Measuring Theory & Precis Instr, Hefei 230009, Peoples R China  
Institution:

### 142. MAUNet: Polyp segmentation network based on multiscale feature fusion of

摘要: Colorectal cancer is a prevalent malignant tumor affecting the digestive tract. Although colonoscopy remains the most effective method for colon examination, it may occasionally fail to detect polyps. In an effort to enhance the detection rate of intestinal polyps during colonoscopy, we propose MAUNet, a polyp segmentation network based on a multi-scale feature fusion of an Attention U-shaped network structure. Our model incorporates advanced components, including the Receptive Field Block, Reverse Attention Block, and Residual Refinement Module, mirroring the analytical process performed by medical imaging professionals. We evaluated the performance of MAUNet on five challenging datasets and conducted a comparative analysis against five state-of-the-art models using six evaluation metrics. The experimental results demonstrate that MAUNet achieves varying levels of performance improvement across the five datasets. Particularly on the Kvasir dataset, the Mean Dice and Mean IOU metrics reached 91.6% and 84.3%, respectively, confirming the model's outstanding performance in polyp segmentation.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0899-9457  
eISSN:  
1098-1098  
入藏号:  
WOS:001258452300001  
Indexed Date:  
2024-07-08  
Institution Address:  
Chongqing Univ Technol, Coll Comp Sci & Engn, Chongqing, Peoples R China  
Institution:

### 143. Context-aware cross-level attention fusion network for infrared small

摘要: Infrared small target detection (IRSTD) plays an essential role in many fields such as air guidance, tracking, and surveillance. However, due to the tiny sizes of infrared small targets, which are easily confused with background noises and lack clear contours and texture information, how to learn more discriminative small target features while suppressing background noises is still a challenging task. In this paper, a context-aware cross-level attention fusion network for IRSTD is proposed. Specifically, a self-attention-induced global context-aware module obtains multilevel attention feature maps with robust positional relationship modeling. The high-level feature maps with abundant semantic information are then passed through a multiscale feature refinement module to restore the target details and highlight salient features. Feature maps at all levels are fed into a channel and spatial filtering module to compress redundant information and remove background noises, which are then used for cross-level feature fusion. Furthermore, to overcome the lack of publicly available datasets, a large-scale multiscene infrared small target dataset with high-quality annotations is constructed. Finally, extensive experiments on both public and our self-developed datasets demonstrate the effectiveness of the proposed method and the superiority compared with other state-of-the-art approaches. (c) 2022 Society of Photo-Optical Instrumentation Engineers  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
1931-3195  
入藏号:  
WOS:000917345800036  
Indexed Date:  
2023-02-17  
Institution Address:  
Zhongyuan Univ Technol, Sch Elect & Informat Engn, Zhengzhou, Peoples R China  
Institution:

### 144. Multiscale Infrared Target Detection Based on Attention Mechanism

摘要: To address the problems of poor textural detail, low contrast, and poor target detection in infrared images, a multiscale infrared target detection model that integrates a channel attention mechanism is proposed based on Yolov4 (You Only Look Once version 4). First, the number of model parameters is reduced by reducing the depth of the backbone feature extraction network. Second, to supplement the shallow highresolution feature information, the multiscale feature fusion module is reconstructed to improve the utilization of the feature information. Finally, before the multiscale feature map is generated, the channel attention mechanism is integrated to further improve the infrared feature extraction ability and reduce noise interference. The experimental results show that the size of the algorithm model in this study was only 28.87% of the Yolov4. The detection accuracy of the infrared targets also significantly improved.  
摘要:  
针对红外图像存在细节纹理特征差、对比度低、目标检测效果差等问题,基于YOLOv4(You Only Look Once version 4)架构提出了一种融合通道注意力机制的多尺度红外目标检测模型。该模型首先通过降低主干特征提取网络深度,减少了模型参数。其次,为补充浅层高分辨率特征信息,重新构建多尺度特征融合模块,提高了特征信息利用率。最后在多尺度加强特征图输出前,融入通道注意力机制,进一步提高红外特征提取能力,降低噪声干扰。实验结果表明,本文算法模型大小仅为YOLOv4的28.87%,对红外目标的检测精度得到了明显提升。  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1001-8891  
入藏号:  
CSCD:7510876  
Indexed Date:  
2023-10-28  
Institution Address:  
河北经贸大学信息技术学院, 石家庄, 河北 050061, 中国  
Institution Address:  
College of Information Technology, Hebei University of Economics and Business, Shijiazhuang, Hebei 050061, China  
Institution:  
河北经贸大学信息技术学院  
Institution:  
College of Information Technology, Hebei University of Economics and Business

### 145. ELMGAN: A GAN-based efficient lightweight multi-scale-feature-fusion

摘要: Cell segmentation and counting is a time-consuming and important experimental step in traditional biomedical research. Many current counting methods are Point-based methods which require exact cell locations. However, there are few such cell datasets with detailed object coordinates. Most existing cell datasets only have the total number of cells and a global segmentation annotation. To effectively use existing datasets, we divide the cell counting task into the cell's number prediction and cell segmentation. We propose a GAN-based efficient lightweight multi-scale-feature-fusion multi-task model (ELMGAN). To coordinate the learning of these two tasks, we propose a Norm-Combined Hybrid loss function (NH loss) and use the method of the generative adversarial network to train our networks. We propose a new Fold Beyond-nearest Upsampling method (FBU) in our lightweight and fast multi-scale-feature-fusion multi-task generator (LFMMG), which is twice as fast as the traditional interpolation upsampling method. We use multi-scale feature fusion technology to improve the quality of segmentation images. LFMMG reduces the number of parameters by nearly 50% compared with U-Net and gets better performance on cell segmentation. Compared with the traditional GAN model, our method improves the speed of image processing by nearly ten times. In addition, we also propose a Coordinated Multitasking Training Discriminator (CMTD) to refine the accuracy of the details of the features. Our method achieves non-Point-based counting that no longer needs to annotate the exact position of each cell in the image during the training and achieves excellent results in cell counting and segmentation. (c) 2022 Elsevier B.V. All rights reserved.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
7  
ISSN:  
0950-7051  
eISSN:  
1872-7409  
入藏号:  
WOS:001004270000001  
Indexed Date:  
2023-06-24  
Institution Address:  
Univ Leicester, Sch Comp & Math Sci, Leicester, Leics, England  
Institution:

### 146. Robust zero-watermarking algorithm for diffusion-weighted images based

摘要: Since diffusion-weighted imaging (DWI) images are high-dimensional medical images with rich texture features, existing traditional zero-watermarking algorithms cannot effectively protect the copyright of DWI images while satisfying the conditions of equilibrium and distinguishability. In this paper, a robust zero-watermarking algorithm for DWI images based on multiscale feature fusion is proposed to achieve effective copyright protection of DWI images. Firstly, a Siamese network for multiscale feature extraction based on the imaging characteristics of DWI images is proposed to extract robust features of two-dimensional slice sequences in each diffusion gradient direction by using different transforms. Next, to address the low distinguishability of extracted features due to the similar anatomical structure of different people under the same medical imaging mode, a variety of features based on texture and spatial correlations are integrated as a prior knowledge features into the network to enhance the distinguishability of zero-watermarking. Then, the dimensionality of the extracted features is reduced by using a clustering algorithm, and the key slice images are selected in each category of each diffusion gradient direction by using the clustering results. The eigenvalues of the eigenmatrix of the key slices are extracted by singular value decomposition to generate a binary feature map according to the gray average of the eigenvalue matrix. Finally, the binary feature map and the watermarking signals are modulated by logic chaos, and the zero-watermarking information is generated by performing the XOR operation. The experimental results show that the proposed algorithm is robust against various intentional or unintentional attacks on medical image processing, such as noise, filtering, JPEG compression and geometric attacks. The extraction accuracy rate of watermarking is above 97%, and it is more robust than traditional robust zero-watermarking methods.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0942-4962  
eISSN:  
1432-1882  
入藏号:  
WOS:001019787900001  
Indexed Date:  
2023-07-12  
Institution Address:  
Guizhou Univ, Coll Comp Sci & Technol, State Key Lab Publ Big Data, Guiyang 550025, Peoples R China  
Institution:

### 147. Anchor-free infrared pedestrian detection based on cross-scale feature

摘要: An anchor-free infrared pedestrian detection algorithm is proposed to address the problems of multiscale, partial occlusion and environmental interference in infrared images, which make it difficult for traditional algorithms to perform accurate detection. First, a cross-scale feature fusion module is devised to improve the detection per-formance of multiscale and partially occluded objects. This module fuses the global features of different residual layers and multiscale local region features, thereby improving the multiscale feature fusion capability of the model while expanding the feature receptive field of the backbone network. Second, the object texture features of infrared images are weak owing to the imaging mechanism, which affects the accuracy of detection model. Thus, we construct a hierarchical attention mapping module to enhance the significance of pedestrian features in complex environments while suppressing the background information. Moreover, to improve the detection performance further, a novel dual-branch head detector is designed to decouple the input features from the channel dimension, which can be used for classification and location. Finally, by introducing the anchor-free concept into the network prediction mechanism, the prediction box can learn ground truth regression parame-ters by itself, thereby simplifying the network structure and improving the model generalization ability. The experimental results demonstrate that the proposed algorithm can accurately detect multiscale infrared pedes-trian objects in complex environments and exhibits better objective evaluation indices compared to other pedestrian detection algorithms. The average precision and recall reached 98.78% and 98.67%, respectively.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
14  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
15  
ISSN:  
1350-4495  
eISSN:  
1879-0275  
入藏号:  
WOS:000965760500001  
Indexed Date:  
2023-04-26  
Institution Address:  
Xian Univ Sci & Technol, Coll Elect & Control Engn, Xian 710054, Peoples R China  
Institution:

### 148. A Lightweight Fusion Strategy With Enhanced Interlayer Feature

摘要: Detecting small objects in drone imagery is challenging due to low resolution and background blending, leading to limited feature information. Multiscale feature fusion can enhance detection by capturing information at different scales, but traditional strategies fall short. Simple concatenation or addition operations do not fully utilize multiscale fusion advantages, resulting in insufficient correlation between features. This inadequacy hinders the detection of small objects, especially in complex backgrounds and densely populated areas. To address this issue and efficiently utilize the limited computational resources, we propose a lightweight fusion strategy based on enhanced interlayer feature correlation (EFC) to replace the traditional feature fusion strategy in feature pyramid network (FPN). The semantic expressions of different layers in the feature pyramid are inconsistent. In EFC, the grouped feature focus unit (GFF) enhances the feature correlation of each layer by focusing on the contextual information of different features. The multilevel feature reconstruction module (MFR) effectively reconstructs and transforms the strength and weakness information of each layer in the pyramid to reduce redundant feature fusion and retain more information about small targets in deep networks. It is noteworthy that the proposed method is plug-and-play and can be widely applied to various base networks. Extensive experiments and comprehensive evaluations on VisDrone, unmanned aerial vehicle benchmark object detection and tracking (UAVDT), and microsoft common objects in context (COCO) demonstrate the effectiveness. Using generalized focal loss (GFL) as the baseline on the VisDrone dataset with a large number of small targets, the proposed method improves the detection mean average precision (mAP) by 1.7%, surpassing many lightweight state-of-the-art methods and significantly reducing the Params and GFLOPs at the neck end. The code will be available at https://github.com/nuliweixiao/EFC.git.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001322218300007  
Indexed Date:  
2024-10-10  
Institution Address:  
Beijing Inst Technol, Beijing 100081, Peoples R China  
Institution Address:  
Minist Educ China, Key Lab Photoelect Imaging Technol & Syst, Beijing 100081, Peoples R China  
Institution Address:  
Beijing Inst Technol, Chongqing Innovat Ctr, Chongqing 401135, Peoples R China  
Institution Address:  
Space Engn Univ, Natl Key Lab Space Target Awareness, Beijing 101416, Peoples R China  
Institution:  
Institution:  
Minist Educ China  
Institution:  
Institution:

### 149. Image Smoothing via Multiscale Global Perception

摘要: Image smoothing provides a fundamental operation for image processing, with a broad spectrum of applications. It is a challenging task which requires global analysis on image patterns with scale awareness. Existing deep models for image smoothing are insufficiently efficient in global perception and multi-scale processing. This letter proposes a deep model with an efficient multi-scale fusion architecture and a series of global processing blocks. The architecture enhances multi-scale feature flow by incorporating features of different scales into both the encoder and decoder blocks of a U-shape network, with multi-scale feature fusion modules inserted between the encoder and the decoder. The global processing blocks leverage the multi-axis processing mechanism to achieve joint local and global perception. Benefiting from these two key designs, our proposed model enjoys superiority in both smoothing performance and computational complexity, as demonstrated in the experiments on two benchmark datasets.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1070-9908  
eISSN:  
1558-2361  
入藏号:  
WOS:001166847500004  
Indexed Date:  
2024-03-24  
Institution Address:  
South China Univ Technol, Sch Comp Sci & Engn, Guangzhou 510006, Peoples R China  
Institution Address:  
Inst Super Robot, Guangzhou 510641, Peoples R China  
Institution:  
Institution:  
Inst Super Robot

### 150. ASFF-YOLOv5: Multielement Detection Method for Road Traffic in UAV

摘要: Road traffic elements are important components of roads and the main elements of structuring basic traffic geographic information databases. However, the following problems still exist in the detection and recognition of road traffic elements: dense elements, poor detection effect of multi-scale objects, and small objects being easily affected by occlusion factors. Therefore, an adaptive spatial feature fusion (ASFF) YOLOv5 network (ASFF-YOLOv5) was proposed for the automatic recognition and detection of multiple multiscale road traffic elements. First, the K-means++ algorithm was used to make clustering statistics on the range of multiscale road traffic elements, and the size of the candidate box suitable for the dataset was obtained. Then, a spatial pyramid pooling fast (SPPF) structure was used to improve the classification accuracy and speed while achieving richer feature information extraction. An ASFF strategy based on a receptive field block (RFB) was proposed to improve the feature scale invariance and enhance the detection effect of small objects. Finally, the experimental effect was evaluated by calculating the mean average precision (mAP). Experimental results showed that the mAP value of the proposed method was 93.1%, which is 19.2% higher than that of the original YOLOv5 model.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
7  
Web of Science 核心合集中的 "被引频次":  
38  
被引频次合计:  
40  
eISSN:  
2072-4292  
入藏号:  
WOS:000833345300001  
Indexed Date:  
2022-08-06  
Institution Address:  
Kunming Univ Sci & Technol, Fac Land Resources Engn, Kunming 650093, Yunnan, Peoples R China  
Institution Address:  
Surveying & Mapping Geoinformat Technol Res Ctr P, Kunming 650093, Yunnan, Peoples R China  
Institution:  
Institution:  
Surveying & Mapping Geoinformat Technol Res Ctr P

### 151. A lightweight and efficient one-stage detection framework?

摘要: Deploying high-performance one-stage object detectors on resource-constrained applications is a challenging task. This paper analyzes factors affecting the computational complexity of onestage detectors and proposes a lightweight and efficient detection framework named LEYOLO. Under this framework, a series of efficient feature extraction modules and a novel channel attention module are designed to compose a lightweight backbone network for detection task. To efficiently combine the features extracted from the backbone, a lightweight multiscale feature fusion structure with a weighted fusion method is proposed to avoid the overhead of dimensionality reduction and downsampling. Further, two detectors (i.e., LEYOLOs and LEYOLOm) are developed based on this framework. Experimental results show that LEYOLO achieves state-of-the-art trade-offs between performance and complexity, given only small computational budgets.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
ISSN:  
0045-7906  
eISSN:  
1879-0755  
入藏号:  
WOS:000898785500004  
Indexed Date:  
2023-01-08  
Institution Address:  
Beijing Forestry Univ, Sch Informat Sci & Technol, Beijing 100083, Peoples R China  
Institution Address:  
Natl Forestry & Grassland Adm, Engn Res Ctr Forestry Oriented Intelligent Informa, Beijing 100083, Peoples R China  
Institution:  
Institution:  
Natl Forestry & Grassland Adm

### 152. MAD-YOLO: A quantitative detection algorithm for dense small-scale

摘要: Marine biological resources are abundant, and the reasonable development, research and protection of marine biological resources are of great significance to marine ecological health and economic development. At present, underwater object quantitative detection plays a very important role in marine biological science research, marine species richness survey, and rare species conservation. However, the problems of a large amount of noise in the underwater environment, small object scale, dense biological distribution, and occlusion all increase the detection difficulty. In this paper, a detection algorithm MAD-YOLO (Multiscale Feature Extraction and Attention Feature Fusion Reinforced YOLO for Marine Benthos Detection) is proposed, which is based on improved YOLOv5 is proposed to solve the above problems. To improve the adaptability of the network to the underwater environment, VOVDarkNet is designed as the feature extraction backbone. It uses the intermediate features with different receptive fields to reinforce the ability to extract feature. AFC-PAN is proposed as the feature fusion network so that the network can learn correct feature information and location information of objects at various scales, improving the network's ability to perceive small objects. Label assignment strategy SimOTA and decoupled head are introduced to help the model better handles occlusion and dense distribution problems. Experiments show the MAD-YOLO algorithm increases mAP0.5:0.95 on the URPC2020 dataset from 49.8% to 53.4% compared to the original YOLOv5. Moreover, the advantages of the model are visualized and analyzed by the method of controlling variables in the experimental part. The experiments show that MAD-YOLO is suitable for detecting blurred, dense, and small-scale objects. The model performs well in marine benthos detection tasks and can effectively promote marine life science research and marine engineering implementation. The source code is publicly available at https://github.com/JoeNan1/MAD-YOLO.  
在中国科学引文数据库中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
26  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
14  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
29  
ISSN:  
1574-9541  
eISSN:  
1878-0512  
入藏号:  
WOS:000946784700001  
Indexed Date:  
2023-03-25  
Institution Address:  
Qingdao Univ Sci & Technol, Coll Informat Sci & Technol, Qingdao 266061, Peoples R China  
Institution:

### 153. MFFCI-YOLOv8: A Lightweight Remote Sensing Object Detection Network

摘要: Most current researches primarily focus on improving experimental accuracy using large models, often neglecting the deployment challenges. There is a growing need for lightweight algorithms in certain remote sensing devices. Moreover, remote sensing images (RSIs) often contain numerous small, densely distributed targets, which pose significant detection challenges. To address these issues, we have improved the YOLOv8s network and developed a lightweight remote sensing object detection (RSOD) network based on multiscale features fusion and context information (MFFCI-YOLOv8). This network combines multiscale feature fusion and contextual information to accurately detect objects in RSIs. First, we introduce the lightweight CSP bottleneck with attention module, which utilizes partial convolution calculation and SimAM attention mechanisms to decrease the number of parameters and computational complexity while enhancing feature extraction capabilities. Second, we design the gate spatial pyramid pooling fast module to enhance the model's perception of scale and contextual information, thus improving the detection of small objects. Last, we employ the multiscale fusion lightweight neck module for more efficient multiscale feature fusion, preventing the loss of small objects. Compared to YOLOv8s, our overall model reduces the number of parameters by 7.7% and FLOPs by 11.9%. We validated the accuracy of MFFCI-YOLOv8 on two remote sensing datasets, NWPU VHR-10 and VisDrone. The experimental results demonstrate that our model offers a low computational cost and high detection accuracy compared to other RSOD models and other YOLO models.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001351591100004  
Indexed Date:  
2024-11-16  
Institution Address:  
Zhengzhou Univ Light Ind, Sch Elect Informat, Zhengzhou 450002, Peoples R China  
Institution Address:  
Zhengzhou Univ Light Ind, Sch Comp Sci & Technol, Zhengzhou 450002, Peoples R China  
Institution Address:  
Nanjing Audit Univ, Jiangsu Modern Intelligent Audit Integrated Applic, Sch Comp Sci, Nanjing 211815, Peoples R China  
Institution:  
Institution:  
Institution:

### 154. An Anchor-Free Defect Detector for Complex Background Based on Pixelwise

摘要: Defect detection is a task to locate and classify the possible defects in an image. However, unlike common object detection tasks, defect detection often needs to deal with images with relatively complex backgrounds, for example, in industrial product quality inspection scenario. The complex background can greatly interfere with the feature of the target objects in the multiscale feature fusion process and therefore puts great challenge on the defect detector. In this work, a channel-space adaptive enhancement feature pyramid network (CA-FPN) is proposed to eliminate this interference from the complex background. By extracting the inner relationship of different scale features, CA-FPN realizes adaptive fusion of multiscale features to enhance the semantic information of the defect while avoiding background interference as much as possible. In particular, CA-FPN is very lightweight. Moreover, considering that defects are often of varying sizes and can be extremely tiny or slender, a flexible anchor-free detector CA-AutoAssign is proposed by combining CA-FPN and an anchor-free detection strategy AutoAssign. Based on the Alibaba Cloud Tianchi Fabric dataset and NEU-DET, CA-AutoAssign is compared with the state-of-the-art (SOTA) detectors. The experimental results show that CA-AutoAssign has the best detection performance with AP50 [mean average precision (mAP) with the intersection over union (IOU) threshold of 50%] reaching 89.1 and 82.7, respectively. Despite the improvement in accuracy, the processing time has barely increased. Furthermore, CA-FPN is applied to other classical detectors, and the experimental results demonstrate the competitiveness and generalization ability of CA-FPN. The code is available at https://github.com/EasonLuht/CA-AutoAssign.git.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
13  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
13  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000922147200094  
Indexed Date:  
2023-05-23  
Institution Address:  
Zhejiang Sci Tech Univ, Sch Informat Sci & Technol, Key Lab Intelligent Text & Flexible Interconnect, Hangzhou 310018, Peoples R China  
Institution:

### 155. Adaptive weighted multiscale feature fusion for small drone object

摘要: Drone object detection in low-altitude airspace plays an essential role in many practical applications, such as security and airspace monitoring. Despite the remarkable progress made by many methods, drone object detection still remains challenging due to the complex background and huge differences in scales of drones. To address the above issues, an improved fully convolutional one-stage object detection (FCOS) model based on adaptive weighted feature fusion (AWFF) module is proposed for multiscale drone object detection in complex background. By learning the spatial relevance of feature maps at each scale and improving the scale invariance of features based on the channel attention mechanism, AWFF module could adaptively fuse the features of adjacent scale. In addition, a receptive field enhancement module is designed to reduce the information loss in the feature fusion process. Extensive experiments are conducted to evaluate the effectiveness of the proposed module and method on the constructed low-altitude drone dataset, which concludes that the mean average precision of the AWFF-FCOS is increased by 2.1% compared with the baseline method. And extensive ablation experiments further demonstrate that the proposed AWFF module and REF module could be integrated into the state-of-the-art method to improve the performance of drone object detection. (C) 2022 Society of Photo-Optical Instrumentation Engineers (SPIE)  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
4  
eISSN:  
1931-3195  
入藏号:  
WOS:000867557000019  
Indexed Date:  
2022-11-09  
Institution Address:  
Nanjing Univ Sci & Technol, Sch Energy & Power Engn, Nanjing, Peoples R China  
Institution Address:  
Xian Res Inst High Tech, Xian, Peoples R China  
Institution:  
Institution:

### 156. Contrastive learning-based generative network for single image deraining

摘要: The crux of image deraining stems from the challenge of recognizing the diverse rain patterns within the rainy image. Most methods for image deraining remain visible rain residuals in the restored image, which suffers from insufficient modeling of rain streaks. In this work, we propose contrastive learning-based generative network (CLGNet), which follows a coarse-to-fine framework. In the coarse phase, our CLGNet employs the hierarchical encoder-decoder structure to remove obvious rain patterns, and first generates the coarse background image. Then, we introduce a well-designed multiscale feature aggregation module in the refining phase to extract and integrate global information dependencies from different scales. In additon, to facilitate the intra-stage and cross-stage information interaction, we propose the intra-stage feature fusion module and the cross-stage feature fusion module to encode broad contextual information. More importantly, we propose an innovative contrastive learning strategy and apply it to each stage of our proposed CLGNet to enhance the decoupling ability of the encoder and help the model recognize complex rain patterns. Extensive experiments on five benchmark datasets demonstrate the superiority of our proposed CLGNet than other state-of-the-art methods for single image deraining on both the visual quality and quantitative evaluation. (C) 2022 SPIE and IS&T  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:000789627500022  
Indexed Date:  
2022-05-27  
Institution Address:  
Xinyang Agr & Forestry Univ, Sch Informat Engn, Xinyang, Peoples R China  
Institution Address:  
Wuhan Polytech Univ, Sch Math & Comp Sci, Wuhan, Peoples R China  
Institution:  
Institution:

### 157. Multi-Feature Fusion for Camouflaged Object Segmentation

摘要: In the field of camouflaged object segmentation,how to extract high-resolution semantic features from a depth model is the key to constructing a target segmentation model.In order to better solve this problem,a new camouflage target segmentation method based on multi-level feature fusion is proposed.A multi-stage gate control module is introduced to selectively fuse the multi-stage middle layer features of Res2Net-50,which can effectively filter the interference information of each level feature map during the feature encoding process.And in decoding,the self-interaction residual module has been used to drive the cross-fusion of encoding features of different scales,which guarantees the obtaining of more accurate target representation information.In addition,this paper combines cross entropy loss and Dice loss as a joint loss function to help the model segment the camouflaged target more accurately.Experimental results show that the proposed model performs better than the other eight typical models in the complex background camouflage data set and three common natural camouflage datasets.  
摘要:  
在伪装目标分割任务中,如何提取深度模型下高分辨率的目标语义特征是构建目标分割模型的关键。针对此问题,提出了一种基于多级特征融合的伪装目标分割方法。在特征编码过程中,引入多级门控模块对Res2Net-50的多级中间层特征进行选择性融合,有效过滤各级特征图的干扰信息;在解码过程中,通过自交互残差模块驱动不同尺度的编码特征实现交叉融合,获得更准确的目标表示信息。此外,在交叉熵损失的基础上加入Dice损失形成联合损失函数,帮助模型更精准地分割伪装目标。实验结果证明,在背景复杂的迷彩伪装数据集以及三个常用自然伪装数据集上,相比其他典型模型,该模型表现出更好的分割效果。  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1002-8331  
入藏号:  
CSCD:7303807  
Indexed Date:  
2023-01-07  
Institution Address:  
陆军工程大学指挥控制工程学院, 南京, 江苏 210007, 中国  
Institution Address:  
陆军工程大学指挥控制工程学院;;陆军炮兵防空兵学院南京校区火力系;;安徽省偏振成像与探测重点实验室, ;;;;安徽省偏振成像与探测重点实验室, 南京;;南京;;合肥, ;;;; 210007;;211100;;230031  
Institution Address:  
Institute of Command-and-Control Engineering,Army Engineering University of PLA, Nanjing, Jiangsu 210007, China  
Institution Address:  
Institute of Command-and-Control Engineering,Army Engineering University of PLA;;Firepower Department,The Army Artillery and Defense Academy of PLA;;The Key Laboratory of Polarization Imaging Detection Technology of Anhui Province, ;;;;The Key Laboratory of Polarization Imaging Detection Technology of Anhui Province, Nanjing;;Nanjing;;Hefei, ;;;; 210007;;211100;;230031  
Institution:  
陆军工程大学指挥控制工程学院  
Institution:  
陆军工程大学指挥控制工程学院;;陆军炮兵防空兵学院南京校区火力系;;安徽省偏振成像与探测重点实验室  
Institution:  
Institute of Command-and-Control Engineering,Army Engineering University of PLA  
Institution:  
Institute of Command-and-Control Engineering,Army Engineering University of PLA;;Firepower Department,The Army Artillery and Defense Academy of PLA;;The Key Laboratory of Polarization Imaging Detection Technology of Anhui Province

### 158. Multiscale Feature Fusion Back-projection Network for Image

摘要: Aiming at the problems that existing image super-resolution reconstruction methods have weak ability to restore image high-frequency details and insufficient feature utilization,a multi-scale feature fusion back projection network is proposed for image super-resolution reconstruction.The network first uses multi-scale convolution kernels in the shallow feature extraction layer to extract feature information of different dimensions to enhance crosschannel information fusion;then builds a multi-scale back projection module to perform feature mapping through recursive learning to improve the early reconstruction capabilities of the network;Finally,local residual feedback is combined with global residual learning to promote the spread and utilization of features,thereby fusing feature information of different depths for image reconstruction.The experimental results of $ \times 2 \sim \times 8$ SR on the images show that the quality of SR image of this method is better than the existing image super-resolution method in subjective perception and objective evaluation index,and the reconstruction performance is relatively better when the scale factors is large.  
摘要:  
针对现有图像超分辨率重建方法恢复图像高频细节能力较弱、特征利用率不足的问题,提出了一种多尺度特征融合反投影网络用于图像超分辨率重建.该网络首先在浅层特征提取层使用多尺度的卷积核提取不同维度的特征信息,增强跨通道信息融合能力;然后,构建多尺度反投影模块通过递归学习执行特征映射,提升网络的早期重建能力;最后,将局部残差反馈结合全局残差学习促进特征的传播和利用,从而融合不同深度的特征信息进行图像重建.对图像进行$ \times 2 \sim \times 8$超分辨率的实验结果表明,本方法的重建图像质量在主观感受和客观评价指标上均优于现有图像超分辨率重建方法,超分辨率倍数大时重建性能相比更优秀.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
10  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
15  
ISSN:  
0254-4156  
入藏号:  
CSCD:7022132  
Indexed Date:  
2021-10-22  
Institution Address:  
南京信息工程大学电子与信息工程学院, 南京, 江苏 210044, 中国  
Institution Address:  
南京信息工程大学电子与信息工程学院;;南京信息工程大学, ;;江苏省大气环境与装备技术协同创新中心, 南京;;南京, ;; 210044;;210044  
Institution Address:  
School of Electronics and Information Engineering,Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China  
Institution Address:  
School of Electronics and Information Engineering,Nanjing University of Information Science and Technology;;Nanjing University of Information Science and Technology, ;;Jiangsu Collaborative Innovation Center for Atmospheric Environment and Equipment Technology, Nanjing;;Nanjing, ;; 210044;;210044  
Institution:  
南京信息工程大学电子与信息工程学院  
Institution:  
南京信息工程大学电子与信息工程学院;;南京信息工程大学  
Institution:  
School of Electronics and Information Engineering,Nanjing University of Information Science and Technology  
Institution:  
School of Electronics and Information Engineering,Nanjing University of Information Science and Technology;;Nanjing University of Information Science and Technology

### 159. Object Detection For Remote Sensing Image Based on Multiscale Feature

摘要: Object detection in remote sensing images is a fundamental task in image analysis and interpretation. We proposed a Multiscale Dilated Convolution Feature Fusion Detector (MDCF2Det) to achieve precise object detection in remote sensing by addressing the problems of multiscale objects and the complexity of the background. To begin, we improve the original feature pyramid network by replacing the general convolution with the dilated convolution to increase the receptive field. Second, to take full advantage of different levels of semantic and location information, we add a skip connection operation from the input node to the output node. Finally, to suppress the noise and highlight the foreground, we add the multi-dimensional attention model before the regional proposal network, to achieve more accurate object detection in remote sensing images. Experiments are carried out on the DOTA and RSOD datasets, and the proposed algorithm's mean average precision reaches 92.95% and 73.39% respectively. The results show that the proposed algorithm can significantly improve the object detection accuracy of remote sensing images.  
摘要:  
在图像分析与解译中,遥感影像目标检测是一项基础性的工作。针对遥感影像目标尺度多样和背景复杂等问题,提出了一种多尺度空洞卷积特征融合检测器(MDCF2Det)来实现遥感目标的精确检测。首先,改进原始特征金字塔网络(FPN),用空洞卷积代替普通卷积,增大感受野;其次,增加从输入节点到输出节点的跳跃连接操作以充分地利用不同层级的语义和位置信息;最后,为了抑制噪声并突出前景,在区域候选网络前增加多维注意力机制模块,从而实现更精确的遥感影像目标检测。在DOTA和RSOD数据集上进行了实验,所提算法的mean average precision(mAP)分别达到了92.95%和73.39%。实验结果表明,所提算法能够有效提升遥感影像目标检测精度。  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7287956  
Indexed Date:  
2023-09-14  
Institution Address:  
兰州交通大学测绘与地理信息学院;;地理国情监测技术应用国家地方联合工程研究中心;;甘肃省地理国情监测工程实验室, ;;地理国情监测技术应用国家地方联合工程研究中心;;甘肃省地理国情监测工程实验室, 兰州;;兰州;;兰州, 甘肃;;甘肃;;甘肃 730070;;730070;;730070, 中国  
Institution Address:  
Faculty of Geomatics, Lanzhou Jiaotong University;;National-Local Joint Engineering Research Center of Technologies and Applications for National Geographic State Monitoring;;Gansu Provincial Engineering Laboratory for National Geographic State Monitoring, ;;National-Local Joint Engineering Research Center of Technologies and Applications for National Geographic State Monitoring;;Gansu Provincial Engineering Laboratory for National Geographic State Monitoring, Lanzhou;;Lanzhou;;Lanzhou, Gansu;;Gansu;;Gansu 730070;;730070;;730070  
Institution:  
兰州交通大学测绘与地理信息学院;;地理国情监测技术应用国家地方联合工程研究中心;;甘肃省地理国情监测工程实验室  
Institution:  
Faculty of Geomatics, Lanzhou Jiaotong University;;National-Local Joint Engineering Research Center of Technologies and Applications for National Geographic State Monitoring;;Gansu Provincial Engineering Laboratory for National Geographic State Monitoring

### 160. MFO-Net: A Multiscale Feature Optimization Network for UAV Image Object

摘要: Object detection in scenes captured by unmanned aerial vehicles (UAVs) is an active research area. However, the performance and efficiency of current small object detection models for UAV images are far from reaching the desired level. The inherent limitations of the features of the small objects themselves and the inconsistency of the contextual information in the feature maps lead to a degradation of the final detection performance. In this letter, to improve the performance of UAV image small object detection, we propose a multiscale feature optimization network, named MFO-Net. We have designed three crucial modules: feature optimization fusion (FOF) module, multiscale localized feature aggregation (MLFA) module, and feature enhancement (FE) module. FOF module enhances the fusion of features with inconsistent contexts at different levels by learning pixelwise displacement, facilitating more effective feature fusion, which further helps focus on and capture critical information about small objects. MLFA module aggregates richer contextual information through multibranch stripe convolution blocks, while the FE module extracts richer gradient flow information, suppresses incompatible information, and enhances feature representation capability. We conduct extensive experiments on the challenging VisDrone2019 dataset and compare the results against those obtained from the state-of-the-art methods. The experimental results show that MFO-Net performs better than other detectors. Specifically, MFO-Net achieves the best performance with 22.3% AP, 38.9% AP(50) , and 22.5% AP(75) on VisDrone2019.  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001197826200008  
Indexed Date:  
2024-04-18  
Institution Address:  
Fujian Agr & Forestry Univ, Coll Comp & Informat Sci, Fuzhou 350002, Peoples R China  
Institution Address:  
Minjiang Univ, Coll Comp & Control Engn, Fuzhou 350108, Peoples R China  
Institution Address:  
Fujian Prov Univ, Fujian Agr & Forestry Univ, Key Lab Smart Agr & Forestry, Fuzhou 350002, Peoples R China  
Institution:  
Institution:  
Institution:

### 161. Wafer defect classification network with global and local multi-scale

摘要: In semiconductor manufacturing, wafer defect classification is an important step in ensuring product quality. However, due to the diversity and complexity of wafer defects, the existing hybrid wafer defect classification network still has shortcomings in accuracy. To solve this problem, a hybrid wafer defect classification network based on global and local multi-scale feature fusionMLG-Net was proposed. MLG-Net consists of three main modules: feature extraction module, global branch, and local branch. The network aims to better extract and utilize the global semantic information and local detail features of wafer defect images, which are combined with multiscale feature fusion technology to form a more comprehensive feature representation, which helps the classifier to make more accurate judgments in the face of complex mixed defects, thereby improving the classification accuracy. To verify the effectiveness of MLG-Net, a large number of experiments were carried out on MixedWM38, a dataset containing 38 mixed types of defects, and the classification accuracy reached 98.84%. The results show that MLG-Net is superior to the six mainstream wafer defect classification methods in terms of comprehensive performance. This result demonstrates the importance and effectiveness of global and local feature fusion in dealing with hybrid wafer defect classification tasks.  
摘要:  
在半导体制造领域, 晶圆缺陷分类是确保产品质量的重要步骤。然而, 由于晶圆缺陷的多样性和复杂性, 现有的混合型晶圆缺陷分类网络在准确性上仍然存在不足。针对这一问题, 提出了一种基于全局和局部多尺度特征融合的混合型晶圆缺陷分类网络MLG-Net。MLG-Net由3个主要模块组成:特征提取模块、全局分支和局部分支。该网络旨在更好地提取和利用晶圆缺陷图像的全局语义信息与局部细节特征, 这两种特征通过多尺度特征融合技术相结合, 最终形成一个更加全面的特征表示, 有助于分类器在面对复杂混合缺陷时, 做出更为准确的判断, 从而提升分类精度。为了验证MLG-Net的有效性, 在包含38种混合类型缺陷的数据集MixedWM38上进行了大量实验, 其分类准确度达到98.84%。结果表明, MLG-Net在综合性能上优于当前主流的六种晶圆缺陷分类方法。这一结果证明了全局与局部特征融合在处理混合型晶圆缺陷分类任务中的重要性和有效性。  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1000-7105  
入藏号:  
CSCD:7860625  
Indexed Date:  
2025-02-15  
Institution Address:  
兰州理工大学电气工程与信息工程学院, 兰州, 甘肃 730000, 中国  
Institution Address:  
College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou, Gansu 730000, China  
Institution:  
兰州理工大学电气工程与信息工程学院  
Institution:  
College of Electrical and Information Engineering, Lanzhou University of Technology

### 162. Hyperspectral Image Classification Based on Multi-Scale Feature Fusion

摘要: In order to fully extract the spectral-spatial features of hyperspectral image (HSI) and to achieve high-precision ground object classification of HSI, an end-to-end multi-scale feature fusion identity (MFFI) block is proposed. This block combines 3D multi-scale convolution, feature fusion and residual connection. Through this block, multi-scale spectralspatial joint features of HSI can be extracted. Because of the end-to-end feature of the block, the final MFFI network can he obtained by stacking multiple MFFI blocks. The average overall accuracy of 99. 73% average accuracy of 99. 84% and Kappa coefficient of 0. 9971 are obtained on three HSI datasets: Salinas, Indian Pines and University of Pavia. The results show that the proposed MFFI block can effectively extract the spectral-spatial features of different types of ground object datasets and achieve satisfactory classification results.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
1006-4125  
入藏号:  
WOS:000892527800015  
Indexed Date:  
2022-12-21  
Institution Address:  
Guizhou Univ, Coll Big Data & Informat Engn, Engn Res Ctr Semicond Power Device Reliabil, Minist Educ, Guiyang 550025, Guizhou, Peoples R China  
Institution Address:  
Chinese Acad Sci, Suzhou Inst Nanotech & Nanobion, Suzhou 215123, Jiangsu, Peoples R China  
Institution:  
Institution:

### 163. Re-Parameterized YOLOv8 Pavement Disease Detection Algorithm

摘要: Road disease detection is an important way to ensure people's traffic safety. In order to improve the accuracy of road disease detection and achieve timely and accurate road disease detection, a pavement disease detection model of re-parameterized YOLOv8 is proposed. First of all, CNX2f module is introduced into the backbone network to improve the ability of the network to extract pavement disease features, and effectively solve the problem that the pavement disease features are easily confused with the background environmental features. Secondly, RepConv and DBB reparameterization modules are introduced to enhance the capability of multi-scale feature fusion and solve the problem of large scale difference of pavement diseases. At the same time, the shared parameter structure of the head is improved, and RBB reparameterization module is introduced to solve the problem of head parameter redundancy and improve the feature extraction capability. Finally, the SPPF\_Avg module is introduced to solve the problem of pavement feature loss and enrich the multi-scale feature expression. The experimental results show that the accuracy of the improved road disease detection network is 73.3%, the recall rate is 62.3% and the mAP is 69.3%, which is 2.6, 3.0 and 2.8 percentage points higher than that of the YOLOv8 network, and the detection effect of the model is improved.  
摘要:  
路面病害检测是保障人民交通安全的重要方式,为了提高路面病害检测的准确率,实现及时、精准的路面病害检测,提出了一种重参数化YOLOv8路面病害检测算法。在主干网络引入CNX2f模块,提高网络对路面病害特征的提取能力,有效解决路面病害特征与背景环境特征易混淆问题;引入RepConv和DBB重参数化模块,增强多尺度特征融合能力,解决路面病害尺度差异较大问题;改进头部采用共享参数结构,并引入RBB重参数模块,解决头部参数冗余问题,并提高特征提取能力;引入SPPF\_Avg模块,解决路面病害特征丢失问题,丰富多尺度特征表达。实验结果表明,改进后的路面病害检测网络精度为73.3%、召回率为62.3%、mAP为69.3%,较YOLOv8网络分别提高了2.6、3.0、2.8个百分点,提高了模型的检测效果。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1002-8331  
入藏号:  
CSCD:7679906  
Indexed Date:  
2024-05-31  
Institution Address:  
华北理工大学电气工程学院, 唐山, 河北 063000, 中国  
Institution Address:  
School of Electrical Engineering,North China University of Science and Technology, Tangshan, Hebei 063000, China  
Institution:  
华北理工大学电气工程学院  
Institution:  
School of Electrical Engineering,North China University of Science and Technology

### 164. Multiscale Feature Fusion and Semi-Supervised Temporal-Spatial Learning

摘要: This article studies the performance monitoring problem for the potassium chloride flotation process, which is a critical component of potassium fertilizer processing. To address its froth image segmentation problem, this article proposes a multiscale feature extraction and fusion network (MsFEFNet) to overcome the multiscale and weak edge characteristics of potassium chloride flotation froth images. MsFEFNet performs simultaneous feature extraction at multiple image scales and automatically learns spatial information of interest at each scale to achieve efficient multiscale information fusion. In addition, the potassium chloride flotation process is a multistage dynamic process with massive unlabeled data. To overcome its dynamic time-varying and working condition spatial similarity characteristics, a semi-supervised froth-grade prediction model based on a temporal-spatial neighborhood learning network combined with Mean Teacher (MT-TSNLNet) is proposed. MT-TSNLNet designs a new objective function for learning the temporal-spatial neighborhood structure of data. The introduction of Mean Teacher can further utilize unlabeled data to promote the proposed prediction model to better track the concentrate grade. To verify the effectiveness of the proposed MsFEFNet and MT-TSNLNet, froth image segmentation and grade prediction experiments are performed on a real-world potassium chloride flotation process dataset.  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
40  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
40  
ISSN:  
2168-2267  
eISSN:  
2168-2275  
入藏号:  
WOS:001043284500001  
PubMed ID:  
37535488  
Indexed Date:  
2023-08-20  
Institution Address:  
Cent South Univ, Sch Automat, Changsha 410083, Peoples R China  
Institution:

### 165. Identification of apple leaf diseases based on MA-ConvNext network and

摘要: The backgrounds are cluttered, the spot sizes of apple leaf disease are varying in complex environments, and the existing models have the problems of multiple parameters and a large amount of calculation. Thus, an apple leaf disease recognition network, ConvNext network based on attention and multiscale feature fusion (MAConvNext), was proposed. A multiscale spatial reconstruction and channel reconstruction block (MSCB) and a feature extraction block with triplet attention fusion (TAFB) were utilized to effectively extract the features at different scales and enhance the focus on leaf disease spots. Additionally, a stepwise relational knowledge distillation method was employed to fuse the "teacher" network (MA-ConvNext) with an "intermediate" network (DenseNet121) to guide the training of the "student" network (EfficientNet-B0) and achieve the model lightweighting. Experimental results showed that MA-ConvNext achieved a recognition accuracy of 99.38%, improving by 3.98 percentage points, 7.55 percentage points and 4.27 percentage points compared to ResNet50, MobileNet-V3, and EfficientNet-V2 networks, respectively. After the stepwise relational knowledge distillation, the recognition accuracy further improved by 1.76 percentage points, with a smaller network size and parameters of 1.56\*10~7 and 5.29\*10~6. respectively. The proposed method offers new insights and technical support for the precise detection of pests and diseases in agriculture.  
摘要:  
针对复杂环境下苹果叶片病害图像背景杂乱、病斑大小不一,以及现有模型参数多、计算量大的问题,提出基于注意力和多尺度特征融合的苹果叶片病害识别网络(MA-ConvNext).通过引入多尺度空间通道重组块(MSCB)和融合三分支注意力机制的特征提取模块(TAFB),有效提取苹果叶片病害图像不同尺度的特征,增强模型对叶片病斑的关注.采用分步关系知识蒸馏方法,将教师网络(MA-ConvNext)和中间网络(DenseNet121)融合,指导学生网络(EfficientNet-B0)训练,实现模型轻量化.实验结果表明, MA-ConvNext网络识别准确率为99.38%,较ResNet50、MobileNet-V3和EfficientNet-V2网络分别提高了3.98个百分点、7.55个百分点和4.27个百分点.经过分步关系知识蒸馏后,识别准确率较蒸馏前提高了1.76个百分点,并且具有更小的网络规模和参数量,分别为1.56\*10~7、5.29\*10~6.所提方法能为后续精准农业的病虫害检测提供新思路和技术支持.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1008-973X  
入藏号:  
CSCD:7804929  
Indexed Date:  
2024-11-23  
Institution Address:  
西安工程大学电子信息学院, 西安, 陕西 710048, 中国  
Institution Address:  
山西大学生命科学学院, 太原, 山西 030031, 中国  
Institution Address:  
School of Electronics and Information, Xian Polytechnic University, Xian, 710048  
Institution Address:  
School of Life Science, Shanxi University, Taiyuan, Shanxi 030031, China  
Institution:  
西安工程大学电子信息学院  
Institution:  
山西大学生命科学学院  
Institution:  
School of Electronics and Information, Xian Polytechnic University  
Institution:  
School of Life Science, Shanxi University

### 166. Semantic Segmentation for Road Scene Based on Multiscale Feature Fusion

摘要: A lightweight network model based on multiscale feature information fusion (MIFNet) is developed in this study owing to the imbalance among the parameter amount, inference speed, and accuracy in many existing semantic segmentation network models. The MIFNet is constructed on the encoding-decoding architecture. In the encoding part, the split strategy and asymmetric convolution are flexibly applied to design lightweight bottleneck structure for feature extraction. The spatial attention mechanism and Laplace edge detection operator are introduced to fuse spatial and edge information to obtain rich feature information. In the decoding part, a new decoder is designed by introducing a channel attention mechanism to recover the size and detail information of the feature map for a complete semantic segmentation task. The MIFNet achieves accuracies of 73.1% and 67.7% on the Cityscapes and CamVid test sets, respectively, with only approximately 0.82 M parameters. Correspondingly, it reaches up to 73.68 frame/s and 85.16 frame/s inference speed, respectively using a single GTX 1080Ti GPU. The results show that the method achieves a good balance in terms of the parameter amount, inference speed, and accuracy, yielding a lightweight, fast, and accurate semantic segmentation.  
摘要:  
针对现有语义分割网络模型难以在参数量、推理速度和精确度中取得平衡的问题,设计了一种多尺度特征信息融合的轻量级网络模型(MIFNet)。MIFNet采用编码-解码结构,在编码部分利用分离策略和非对称卷积设计了轻量型特征提取瓶颈结构,且引入空间注意力机制与Laplace边缘检测算子组成边缘-空间融合模块,将空间信息和边缘信息进行融合得到丰富的特征信息。在解码部分引入通道注意力机制恢复特征图尺寸和细节信息完成语义分割。在Cityscapes和CamVid测试集上,MIFNet仅以0.82 M的参数量分别取得了73.1%和67.7%的分割精度,同时在单个GTX 1080Ti GPU下分别获得73.68 frame/s和85.16 frame/s的推理速度,表明该方法在参数量、推理速度和精确度3个指标上得到较好平衡,实现了轻量、快速、精准的语义分割。  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7512309  
Indexed Date:  
2023-10-28  
Institution Address:  
暨南大学信息科学技术学院, 广州, 广东 510632, 中国  
Institution Address:  
College of Information Science and Technology, Jinan University (Guangdong), Guangzhou, Guangdong 510632, China  
Institution:  
暨南大学信息科学技术学院  
Institution:  
College of Information Science and Technology, Jinan University (Guangdong)

### 167. Transmission line defect detection method based on super-resolution

摘要: Aiming at the problem that the quality of the captured image may be poor in the inspection of transmission lines, and the problem that the detection accuracy of traditional methods is not high due to the line defect that the targets are small and densely distributed, a transmission line defect detection method based on super-resolution reconstruction and multi-scale feature fusion is proposed. First, the super-resolution network is used to reconstruct the inspection image, improve the clarity and enrich the feature information contained in the image. Then the improved YOLOX network is used to detect defects in the inspection image, and the convolution block attention mechanism is embedded in the backbone network to strengthen the positioning ability of the model for overlapping small targets. In order to further improve the detection ability of small targets, a shallow detection scale is added to YOLOX's feature fusion network for feature fusion. Finally, by using CIOU to optimize the loss function of the bounding box, improve the convergence ability of the model and reduce the missed detection rate of the defect targets. The experimental results show that the proposed method can accurately detect the transmission line defects on the basis of improving the inspection image quality, with an accuracy of 93.27%. Compared with classical models such as SSD, it has stronger extraction ability and robustness for small and dense defect targets.  
摘要:  
针对输电线路巡检中可能存在拍摄图像质量不高的问题,以及线路缺陷目标小而分布密集而导致传统方法检测精度不高的问题,提出一种基于超分辨率重建与多尺度特征融合的输电线路缺陷检测方法。首先,使用超分辨率网络对巡检图像进行重建,提升清晰度,丰富图像中包含的特征信息;然后使用改进的YOLOX网络检测巡检图像中的缺陷,在主干网络中嵌入卷积块注意力机制,强化模型对重叠小目标的定位能力;为进一步提升小目标的检测能力,在YOLOX的特征融合网络中新增浅层检测尺度进行特征融合;最后,通过使用CIOU优化边界框损失函数提升模型收敛能力,降低缺陷目标的漏检率。实验结果表明,所提方法能在提升巡检图像质量的基础上对输电线路缺陷准确地检测,精度达到93.27%,相比SSD等经典模型,对小而密集的缺陷目标有着更强的提取能力和鲁棒性。  
在中国科学引文数据库中的被引频次:  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
1000-7105  
入藏号:  
CSCD:7439760  
Indexed Date:  
2023-07-14  
Institution Address:  
带电巡检与智能作业技术国网公司实验室;;智能带电作业技术及装备(机器人)湖南省重点实验室, ;;智能带电作业技术及装备(机器人)湖南省重点实验室, 长沙;;长沙, ;; 421000;;421000  
Institution Address:  
智能带电作业技术及装备(机器人)湖南省重点实验室;;三峡大学电气与新能源学院, 智能带电作业技术及装备(机器人)湖南省重点实验室;;, 长沙;;宜昌, ;; 421000;;443002  
Institution Address:  
Live Inspection and Intelligent Operation Technology State Grid Corporation Laboratory;;Hunan Province Key Laboratory of Intelligent Live Working Technology and Equipment(Robot), ;;Hunan Province Key Laboratory of Intelligent Live Working Technology and Equipment(Robot), Changsha;;Changsha, ;; 421000;;421000  
Institution Address:  
Hunan Province Key Laboratory of Intelligent Live Working Technology and Equipment(Robot);;College of Electrical Engineering and New Energy, China Three Gorges University, Hunan Province Key Laboratory of Intelligent Live Working Technology and Equipment(Robot);;, Changsha;;Yichang, ;; 421000;;443002  
Institution:  
带电巡检与智能作业技术国网公司实验室;;智能带电作业技术及装备(机器人)湖南省重点实验室  
Institution:  
智能带电作业技术及装备(机器人)湖南省重点实验室;;三峡大学电气与新能源学院  
Institution:  
Live Inspection and Intelligent Operation Technology State Grid Corporation Laboratory;;Hunan Province Key Laboratory of Intelligent Live Working Technology and Equipment(Robot)  
Institution:  
Hunan Province Key Laboratory of Intelligent Live Working Technology and Equipment(Robot);;College of Electrical Engineering and New Energy, China Three Gorges University

### 168. Single Image Dehazing Network Based on Double Scale Feature Fusion

摘要: Learning-based dehazing methods perform well with a synthetic dataset but have certain problems,such as residual haze and color distortion,in a real scene.To solve these problems,a single-image dehazing network based on double-scale feature fusion is proposed in this paper.The proposed network consists of two parts:feature extraction and multiscale feature fusion. In the feature extraction stage,the local and global features of the input image are extracted through the residual dense and feature extraction blocks,respectively,using a spatial attention mechanism.In the feature fusion stage,the local and global feature maps are channel weighted using the channel attention mechanism and fused through a convolution operation.Finally,the gated network adaptively combines three feature maps from different depths to restore the high-quality haze-free images.The experimental results show that the Peak Signal-to-Noise Ratio(PSNR)of the proposed network under indoor dataset is 33.04 dB,and the Structural Similarity(SSIM)is 0.983.And the PSNR and SSIM are 1.33 dB and 0.041 higher than that of the GirdDehazeNet network under HAZERD dataset,respectively.In addition,there are 0.34M model parameters,and the Floating Point Operations(FLOPs)of the network occur at a rate of 16.06\*10~9 frames/s,it shows that the proposed network achieves ideal dehazing results for both synthetic and real images with low complexity.  
摘要:  
基于深度学习的图像去雾方法在合成数据集上表现良好,但在真实场景中应用时存在去雾不彻底、颜色失真等问题。提出一种新的单幅图像去雾网络,该网络包含特征提取、特征融合2个模块。在特征提取模块中,通过残差密集块和具有空间注意机制的特征提取块分别提取图像的局部特征和全局特征。在特征融合模块中,利用通道注意力机制对局部特征图和全局特征图进行通道加权,并通过卷积操作融合加权后的局部特征图与全局特征图。最后,采用门控网络自适应结合3个不同深度的融合特征图,以恢复高质量的去雾图像。实验结果表明,所提网络在室内数据集下的峰值信噪比(PSNR)和结构相似度(SSIM)分别为33.04 dB、0.983,在HAZERD数据集下的PSNR和SSIM分别比GridDehazeNet网络高出1.33 dB和0.041。同时,该网络的模型参数量和浮点运算数分别为0.34M和16.06\*10~9 frame/s,具有较低复杂度,对合成图像和真实图像均可取得理想的去雾效果。  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1000-3428  
入藏号:  
CSCD:7339840  
Indexed Date:  
2023-03-10  
Institution Address:  
火箭军工程大学研究生院, 西安, 陕西 710025, 中国  
Institution Address:  
Graduate School, Rocket Force University of Engineering, Xi'an, Shaanxi 710025, China  
Institution:  
火箭军工程大学研究生院  
Institution:  
Graduate School, Rocket Force University of Engineering

### 169. Remote Sensing Semantic Segmentation via Boundary Supervision-Aided

摘要: High spatial resolution (HSR) remote sensing (RS) images inevitably pose the challenge of multiscale transformation, as small objects, such as cars and helicopters (HCs), may occupy only a few pixel points. This incurs a significant hurdle for global context modeling, particularly in backbone networks with large downsampling coefficients. Simple summation or concatenation techniques, such as skip connections, fail to address semantic gaps and even impose negative impacts on multiscale feature fusion. Meanwhile, due to the complexity of foreground objects, the boundary details of HSR RS images are easy to lose in sampling operations. To overcome these challenges, we propose a multiscale channelwise cross attention network (MCCANet) assisted by boundary supervision (BS). Technically, MCCA captures the channel attention (CA) with various scales, which allows dynamic and adaptive feature fusion in a contextual scale-aware manner and focuses on both large and small objects distributed throughout the inputs. Besides, a channel and context strainer (CCS) module is proposed and embedded in MCCA, filtering channels and contexts for the mitigation of intraclass differences. In addition, we apply a BS module to recover boundary contour, avoiding the blurring effect during the construction of contextual information. The refined boundary allows for the effective recognition of surrounding pixels, ensuring a better segmentation performance. Extensive experiments on the instance segmentation in aerial images dataset (iSAID), International Society for Photogrammetry and Remote Sensing (ISPRS) Potsdam, and land-cover domain adaptive (LoveDA) datasets demonstrate that our proposed MCCANet achieves a good balance of high accuracy and efficiency. Code will be available at: https://github.com/ZhengJianwei2/MCCANet.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
16  
被引频次合计:  
16  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001033534200006  
Indexed Date:  
2023-08-19  
Institution Address:  
Zhejiang Univ Technol, Coll Comp Sci & Technol, Hangzhou 310023, Peoples R China  
Institution:

### 170. A YOLOW Algorithm of Water-Crossing Object Detection

摘要: Accurately identifying and locating water-crossing objects is of utmost importance for environmental protection. However, traditional detection algorithms often exhibit poor anti-interference performance and low detection accuracy in complex environments. To address these issues, this paper proposes a YOLOW algorithm based on a one-stage object detection algorithm for the automatic identification of water objects. The proposed algorithm incorporates two new modules, namely the SPDCS module and the SPPAUG module, to improve the model's performance. Specifically, the SPDCS module retains all information in the channel dimension, thereby enhancing the model's detection accuracy and recognition ability for water-crossing objects. The SPPAUG module performs multiscale feature fusion, which further improves the model's detection accuracy and recognition ability. Moreover, the C2f module is introduced from YOLOv8 to increase the detection speed. Experimental results on a water-floating object dataset demonstrate that the improved YOLOW model outperforms the standard YOLOv5s algorithm, especially in water-crossing object detection. This research has significant implications for environmental monitoring and protection.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001045443700001  
Indexed Date:  
2023-08-23  
Institution Address:  
Hohai Univ, Sch Comp & Informat, Nanjing 211100, Peoples R China  
Institution Address:  
Hohai Univ, Key Lab Water Big Data Technol, Minist Water Resources, Nanjing 211100, Peoples R China  
Institution:  
Institution:

### 171. CMCF-Net: An End-to-End Context Multiscale Cross-Fusion Network for

摘要: Image copy-move forgery detection (CMFD) has become a challenging problem due to increasingly powerful editing software that makes forged images increasingly realistic. Existing algorithms that directly connect multiple scales of features in the encoder part may not effectively aggregate contextual information, resulting in poor performance. In this paper, an end-to-end context multiscale cross-fusion network (CMCF-Net) is proposed to detect image copy-move forgery. The proposed network consists of a multiscale feature extraction fusion (MSF) module and a multi-information fusion decoding (MFD) module. Multiscale information is efficiently extracted and fused in the MSF module utilizing stacked-scale feature fusion, which improves the network's forgery localization ability on objects of different scales. The MFD module employs contextual information combination and weighted fusion of multiscale information to guide the network in obtaining relevant clues from correlated information at multiple different scales. Experimental results and analysis have demonstrated that the proposed CMCF-Net achieves the best localization results with higher robustness.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1520-9210  
eISSN:  
1941-0077  
入藏号:  
WOS:001200272600020  
Indexed Date:  
2024-06-11  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Comp Sci, Nanjing 210044, Peoples R China  
Institution Address:  
Minist Educ, Engn Res Ctr Digital Forens, Jiangsu Collaborat Innovat Ctr Atmospher Environm, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Software Engn, Nanjing 210044, Peoples R China  
Institution Address:  
Natl Dong Hwa Univ, Dept Comp Sci & Informat Engn, Hualien 974301, Taiwan  
Institution Address:  
Fudan Univ, Sch Comp Sci, Shanghai 200433, Peoples R China  
Institution:  
Institution:  
Minist Educ  
Institution:  
Institution:  
Institution:

### 172. A coarse-to-fine multi-scale feature hybrid low-dose CT denoising

摘要: With the growing development and wide clinical application of CT technology, the potential radiation damage to patients has sparked public concern. However, reducing the radiation dose may cause large amounts of noise and artifacts in the reconstructed images, which may affect the accuracy of the clinical diagnosis. Therefore, improving the quality of low-dose CT scans has become a popular research topic. Generative adversarial networks (GAN) have provided new research ideas for low-dose CT (LDCT) denoising. However, utilizing only image decomposition or adding new functional subnetworks cannot effectively fuse the same type of features with different scales (or different types of features). Thus, most current GAN-based denoising networks often suffer from low feature utilization and increased network complexity. To address these problems, we propose a coarse-to-fine multiscale feature hybrid low-dose CT denoising network (CMFHGAN). The generator consists of a global denoising module, local texture feature enhancement module, and self-calibration feature fusion module. The three modules complement each other and guarantee overall denoising performance. In addition, to further improve the denoising performance, we propose a multi-resolution inception discriminator with multiscale feature extraction ability. Experiments were performed on the Mayo and Piglet datasets, and the results showed that the proposed method outperformed the state-of-the-art denoising algorithms.  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0923-5965  
eISSN:  
1879-2677  
入藏号:  
WOS:001069608900001  
Indexed Date:  
2023-10-03  
Institution Address:  
Taiyuan Univ Sci & Technol, Taiyuan 030024, Peoples R China  
Institution:

### 173. Object Detection For Remote Sensing Image Based on Multiscale Feature

摘要: Object detection in remote sensing images is a fundamental task in image analysis and interpretation. We proposed a Multiscale Dilated Convolution Feature Fusion Detector (MDCF(2)Det) to achieve precise object detection in remote sensing by addressing the problems of multiscale objects and the complexity of the background. To begin, we improve the original feature pyramid network by replacing the general convolution with the dilated convolution to increase the receptive field. Second, to take full advantage of different levels of semantic and location information, we add a skip connection operation from the input node to the output node. Finally, to suppress the noise and highlight the foreground, we add the multi-dimensional attention model before the regional proposal network, to achieve more accurate object detection in remote sensing images. Experiments are carried out on the DOTA and RSOD datasets, and the proposed algorithm's mean average precision reaches 92. 95% and 73. 39% respectively. The results show that the proposed algorithm can significantly improve the object detection accuracy of remote sensing images.  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
1006-4125  
入藏号:  
WOS:000834824700042  
Indexed Date:  
2022-08-11  
Institution Address:  
Lanzhou Jiaotong Univ, Fac Geomat, Lanzhou 730070, Gansu, Peoples R China  
Institution Address:  
Natl Local Joint Engn Res Ctr Technol & Applicat, Lanzhou 730070, Gansu, Peoples R China  
Institution Address:  
Gansu Prov Engn Lab Natl Geog State Monitoring, Lanzhou 730070, Gansu, Peoples R China  
Institution:  
Institution:  
Natl Local Joint Engn Res Ctr Technol & Applicat  
Institution:  
Gansu Prov Engn Lab Natl Geog State Monitoring

### 174. Dual-branch network object detection algorithm based on dual-modality

摘要: Aiming at the limitations of visible images in object detection, this paper proposes a dual-branch network object detection algorithm based on dual-modality fusion of visible and infrared images. Based on YOLOv7-s, the algorithm firstly introduces a spatial attention module to enhance the model's ability of capturing key features; secondly, to resolve the problem of inconsistent object sizes, a visible multi-scale feature fusion module is proposed, meanwhile, the structure of the SimCSPSPPF module (an improved spatial pyramid pooling module) from YOLOv6 is adopted to construct an infrared multi-scale feature fusion module to efficiently extract multi-scale features from infrared images; finally, a cross-modal feature fusion module is proposed to fuse corresponding scale features from visible and infrared images. The proposed algorithm is tested on KAIST, FLIR, and GIR datasets, experimental results show that the proposed algorithm has better performance, compared with the YOLOv7-s algorithm to detect visible and infrared images separately on the KAIST dataset, the detection accuracy is improved by 18.0 and 5.1%, respectively, and detection speed is 51.8 FPS; on FLIR and GIR datasets, the proposed algorithm also demonstrates significant advantages. Furthermore, the proposed algorithm can detect objects on individual visible or infrared images while maintaining high detection accuracy.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0942-4962  
eISSN:  
1432-1882  
入藏号:  
WOS:001350711200002  
Indexed Date:  
2024-11-15  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Comp Sci & Technol, Xian 710121, Peoples R China  
Institution Address:  
Key Lab Network Data Anal & Intelligent Proc Shaan, Xian 710121, Peoples R China  
Institution Address:  
Air Force Engn Univ, Informat & Nav Coll, Xian 710100, Peoples R China  
Institution:  
Institution:  
Key Lab Network Data Anal & Intelligent Proc Shaan  
Institution:

### 175. MFFRand: Semantic Segmentation of Point Clouds Based on Multi-Scale

摘要: With the application of the random sampling method in the down-sampling of point clouds data, the processing speed of point clouds has been greatly improved. However, the utilization of semantic information is still insufficient. To address this problem, we propose a point cloud semantic segmentation network called MFFRand (Multi-Scale Feature Fusion Based on RandLA-Net). Based on RandLA-Net, a multi-scale feature fusion module is developed, which is stacked by encoder-decoders with different depths. The feature maps extracted by the multi-scale feature fusion module are continuously concatenated and fused. Furthermore, for the network to be trained better, the multi-loss supervision module is proposed, which could strengthen the control of the training process of the local structure by adding sub-losses in the end of different decoder structures. Moreover, the trained MFFRand network could be connected to the inference network by different decoder terminals separately, which could achieve the inference of different depths of the network. Compared to RandLA-Net, MFFRand has improved mIoU on both S3DIS and Semantic3D datasets, reaching 71.1% and 74.8%, respectively. Extensive experimental results on the point cloud dataset demonstrate the effectiveness of our method.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
eISSN:  
2079-9292  
入藏号:  
WOS:000883370300001  
Indexed Date:  
2022-11-25  
Institution Address:  
Shanghai Polytech Univ, Sch Resources & Environm Engn, Shanghai 201209, Peoples R China  
Institution Address:  
Shanghai Polytech Univ, Sch Comp & Informat Engn, Shanghai 201209, Peoples R China  
Institution:  
Institution:

### 176. Recognition of abnormal car door noise based on multi-scale feature

摘要: To accurately identify the abnormal door-closing noise, we propose a method to recognize the time-frequency image of door closing sound based on a multi-scale feature fusion network model. The door-closing sound signal is transformed into a time-frequency image through wavelet analysis, and a classification model based on multi-scale feature fusion is designed. The model introduces multi-scale filters and dilated convolution and adds two improved inception modules to keep the model lightweight. At the same time, richer spatial features can be obtained. The features of different scales are spliced and input to the fully connected layer, and a dropout layer is added to the fully connected layer to suppress overfitting. By comparing the loss and accuracy rate, adjusting different hyperparameters, the optimal model is obtained. The experimental results show that the multi-scale feature fusion network model has a higher accuracy rate than the transfer learning model. Test accuracy rate is 86% and can effectively recognize abnormal door-closing noise. It provides a feasible theoretical basis for the direction of abnormal door noise recognition.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0954-4070  
eISSN:  
2041-2991  
入藏号:  
WOS:000780077600001  
Indexed Date:  
2022-04-22  
Institution Address:  
Shanghai Univ Engn Sci, Sch Mech & Automot Engn, 333 Longteng Rd, Shanghai 201620, Peoples R China  
Institution Address:  
China Automot Technol Res Ctr Co Ltd, Automot Engn Res Inst, Tianjin, Peoples R China  
Institution Address:  
FAW Res & Dev Inst, State Key Lab Comprehens Technol Automobile Vibra, Changchun, Peoples R China  
Institution:  
Institution:  
China Automot Technol Res Ctr Co Ltd  
Institution:  
FAW Res & Dev Inst

### 177. Multi-Stage Multi-Scale Local Feature Fusion for Infrared Small Target

摘要: The detection of small infrared targets with dense distributions and large-scale variations is an extremely challenging problem. This paper proposes a multi-stage, multi-scale local feature fusion method for infrared small target detection to address this problem. The method is based on multi-stage and multi-scale local feature fusion. Firstly, considering the significant variation in target sizes, ResNet-18 is utilized to extract image features at different stages. Then, for each stage, multi-scale feature pyramids are employed to obtain corresponding multi-scale local features. Secondly, to enhance the detection rate of densely distributed targets, the multi-stage and multi-scale features are progressively fused and concatenated to form the final fusion results. Finally, the fusion results are fed into the target detector for detection. The experimental results for the SIRST and MDFA demonstrate that the proposed method effectively improves the performance of infrared small target detection. The proposed method achieved mIoU values of 63.43% and 46.29% on two datasets, along with F-measure values of 77.62% and 63.28%, respectively.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
13  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
13  
eISSN:  
2072-4292  
入藏号:  
WOS:001074241600001  
Indexed Date:  
2023-10-12  
Institution Address:  
Huazhong Univ Sci & Technol, Sch Elect Informat & Commun, Wuhan 430074, Peoples R China  
Institution:

### 178. PFF-CB: Multiscale Occlusion Pedestrian Detection Method Based on PFF

摘要: Occlusion pedestrian detection is an important and difficult task in pedestrian detection. At present, the main method to deal with occlusion pedestrian detection usually adopts pedestrian parts or human body relationship methods. However, in the scene of crowd occlusion or severe pedestrian occlusion, only small parts of the body can be used for detection. Pedestrian parts or human body relationship methods cannot effectively address these issues. In view of the above problems, this paper abandoned the occlusion processing method of pedestrian parts or human body relationship. Considering that it is difficult to establish the relationship between parts and key points. The scale of visible parts of the occlusion pedestrian is small, and the scale of no occlusion pedestrian and occlusion pedestrian in the same picture is different. A multiscale feature attention fusion network named parallel feature fusion with CBAM (PFF-CB) is proposed for occlusion pedestrian detection. Feature information of different scales can be integrated effectively in the PFF-CB module. PFF-CB module uses a convolutional block attention module (CBAM) to enhance the important feature information in space and channel. A parallel feature fusion module based on FPN is used to enhance key features. The performance of the proposed module was tested on two common data sets of occlusion pedestrians with different occlusion types. The results show that the PFF-CB module makes a good performance in occlusion pedestrian detection tasks.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
1687-5265  
eISSN:  
1687-5273  
入藏号:  
WOS:000793377400010  
PubMed ID:  
35498206  
Indexed Date:  
2022-05-23  
Institution Address:  
Shijiazhuang Tiedao Univ, Sch Informat Sci & Technol, Shaoxing, Peoples R China  
Institution Address:  
Hebei Key Lab Electromagnet Environm Effects & In, Shijiazhuang 050043, Hebei, Peoples R China  
Institution:  
Institution:  
Hebei Key Lab Electromagnet Environm Effects & In

### 179. Infrared monocular ranging algorithm based on multiscale feature fusion

摘要: Due to the introduction of MonoDepth2,unsupervised monocular ranging has made great progress in the field of visible light. However,visible light is not applicable in some scenes,such as at night and in some low-visibility environments. Infrared thermal imaging can obtain clear target images at night and under low-visibility conditions,so it is necessary to estimate the depth of infrared image. However,due to the different characteristics of visible and infrared images,it is unreasonable to migrate existing monocular depth estimation algorithms directly to infrared images. An infrared monocular ranging algorithm based on multiscale feature fusion after improving the MonoDepth2 algorithm can solve this problem. A new loss function,edge loss function,was designed for the low texture characteristic of infrared image to reduce pixel mismatch during image reprojection. The previous unsupervised monocular ranging simply upsamples the four-scale depth maps to the original image resolution to calculate projection errors,ignoring the correlation between scales and the contribution differences between different scales. A weighted Bi-directional Feature Pyramid Network(BiFPN)was applied to feature fusion of multiscale depth maps so that the blurring of depth map edge was solved. In addition,Residual Network(ResNet)structure was replaced by Cross Stage Partial Network(CSPNet)to reduce network complexity and increase operation speed. The experimental results show that edge loss is more suitable for infrared image ranging,resulting in better depth map quality. After adding BiFPN structure,the edge of depth image is clearer. After replacing ResNet with CSPNet, the inference speed is improved by about 20 percentage points. The proposed algorithm can accurately estimate the depth of the infrared image,solving the problem of depth estimation in night low-light scenes and some low-visibility scenes,and the application of this algorithm can also reduce the cost of assisted driving to a certain extent.  
摘要:  
由于MonoDepth2的提出,无监督单目测距在可见光领域取得了重大发展;然而在某些场景例如夜间以及一些低能见度的环境,可见光并不适用,而红外热成像可以在夜间和低能见度条件下获得清晰的目标图像,因此对于红外图像的深度估计显得尤为必要。由于可见光和红外图像的特性不同,直接将现有可见光单目深度估计算法迁移到红外图像是不合理的。针对该问题,对MonoDepth2算法进行改进,提出了基于多尺度特征融合的红外单目测距算法。针对红外图像低纹理的特性设计了一项新的损失函数边缘损失函数,旨在降低图像重投影时的像素误匹配。不同于以往的无监督单目测距单纯地将四个尺度的深度图统一上采样到原图像分辨率计算投影误差而忽略了尺度之间的关联性以及不同尺度之间的贡献差异,将加权的双向特征金字塔网络(BiFPN)应用于多尺度深度图的特征融合,解决了深度图边缘模糊问题。另外用跨阶段部分网络(CSPNet)替换残差网络(ResNet)结构,以降低网络复杂度并提高运算速度。实验结果表明,边缘损失更适合红外图像测距,使得深度图质量更高;在加入BiFPN结构之后,深度图像的边缘更加清晰;将ResNet替换为CSPNet之后,推理速度提高了大约20个百分点。该算法能够准确估计出红外图像的深度,解决夜间低光照场景以及一些低能见度场景下的深度估计难题;该算法的应用也可以在一定程度上降低汽车辅助驾驶的成本。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1001-9081  
入藏号:  
CSCD:7181533  
Indexed Date:  
2022-06-10  
Institution Address:  
哈尔滨工程大学信息与通信工程学院;;艾睿光电科技有限公司, ;;, 哈尔滨;;烟台, ;;山东 150001;;264000  
Institution Address:  
哈尔滨工程大学信息与通信工程学院, 哈尔滨, 黑龙江 150001, 中国  
Institution Address:  
艾睿光电科技有限公司, 烟台, 山东 264000, 中国  
Institution Address:  
College of Information and Communication Engineering, Harbin Engineering University;;IRay Technology Company Limited, ;;, Harbin;;Yantai, Heilongjiang;;Shandong 150001;;264000  
Institution Address:  
College of Information and Communication Engineering, Harbin Engineering University, Harbin, Heilongjiang 150001, China  
Institution Address:  
IRay Technology Company Limited, Yantai, Shandong 264000, China  
Institution:  
哈尔滨工程大学信息与通信工程学院;;艾睿光电科技有限公司  
Institution:  
哈尔滨工程大学信息与通信工程学院  
Institution:  
艾睿光电科技有限公司  
Institution:  
College of Information and Communication Engineering, Harbin Engineering University;;IRay Technology Company Limited  
Institution:  
College of Information and Communication Engineering, Harbin Engineering University  
Institution:  
IRay Technology Company Limited

### 180. SAIR-YOLO: An Improved YOLOv8 Network for Sea-Air Background IR

摘要: The performance of IR small-object detection algorithms determines the detection capability and reliability of electrooptic tracking devices in complex environments. For IR images with sea-air backgrounds, reflections on surface waves increase background thermal noise that might occlude or alter small objects, complicating their detection. Additionally, clouds and waves complicate background texture, also reducing object detectability. In this study, we propose a new sea-air background IR (SAIR) detection model on the basis of the YOLOv8 network, called SAIR-YOLO, with three major improvements. First, an asymptotic multiscale feature fusion network gradually integrates different-scale features to mitigate the semantic gap between nonadjacent features while reducing the influence of sea-air background noise on feature representation. Second, a strengthened detection head discriminates irrelevant background features and focuses network attention on the objects. Third, a hybrid intersection-over-union (IoU) loss function improves detection performance, by focusing on shape similarities, and expands the effective regression range. Experimental results yield average SAIR-YOLO precisions of 80.2%, 84.4%, and 96.4% for three distinct datasets: a custom dataset and the SIRST-V2 and NUDT-SIRST publicly available datasets. This represents improvements of 7.0%, 4.9%, and 0.7%, respectively, on the YOLOv8 model.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001410280000018  
Indexed Date:  
2025-02-09  
Institution Address:  
China Univ Geosci, Sch Automat, Wuhan 430074, Peoples R China  
Institution Address:  
Hubei Key Lab Adv Control & Intelligent Automat, Wuhan 430074, Peoples R China  
Institution Address:  
Minist Educ, Engn Res Ctr Intelligent Technol Geoexplorat, Wuhan 430074, Peoples R China  
Institution:  
Institution:  
Hubei Key Lab Adv Control & Intelligent Automat  
Institution:  
Minist Educ

### 181. Revisiting Feature Fusion for RGB-T Salient Object Detection

摘要: While many RGB-based saliency detection algorithms have recently shown the capability of segmenting salient objects from an image, they still suffer from unsatisfactory performance when dealing with complex scenarios, insufficient illumination or occluded appearances. To overcome this problem, this article studies RGB-T saliency detection, where we take advantage of thermal modality's robustness against illumination and occlusion. To achieve this goal, we revisit feature fusion for mining intrinsic RGB-T saliency patterns and propose a novel deep feature fusion network, which consists of the multi-scale, multi-modality, and multi-level feature fusion modules. Specifically, the multi-scale feature fusion module captures rich contexture features from each modality feature, while the multi-modality and multi-level feature fusion modules integrate complementary features from different modality features and different level of features, respectively. To demonstrate the effectiveness of the proposed approach, we conduct comprehensive experiments on the RGB-T saliency detection benchmark. The experimental results demonstrate that our approach outperforms other state-of-the-art methods and the conventional feature fusion modules by a large margin.  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
108  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
109  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:000647394100011  
Indexed Date:  
2021-06-01  
Institution Address:  
Xidian Univ, Minist Educ, Key Lab Elect Equipment Struct Design, Xian 710071, Peoples R China  
Institution Address:  
Xidian Univ, Sch Mechanoelect Engn, Ctr Complex Syst, Xian 710071, Peoples R China  
Institution Address:  
Aberystwyth Univ, Comp Sci Dept, Aberystwyth SY23 3FL, Dyfed, Wales  
Institution:  
Institution:  
Institution:

### 182. Multi-Scale Feature Fusion with Attention Mechanism Based on CGAN

摘要: This paper proposes a colorization algorithm for infrared images based on a Conditional Generative Adversarial Network (CGAN) with multi-scale feature fusion and attention mechanisms, aiming to address issues such as color leakage and unclear semantics in existing infrared image coloring methods. Firstly, we improved the generator of the CGAN network by incorporating a multi-scale feature extraction module into the U-Net architecture to fuse features from different scales, thereby enhancing the network's ability to extract features and improving its semantic understanding, which improves the problems of color leakage and blurriness during colorization. Secondly, we enhanced the discriminator of the CGAN network by introducing an attention mechanism module, which includes channel attention and spatial attention modules, to better distinguish between real and generated images, thereby improving the semantic clarity of the resulting infrared images. Finally, we jointly improved the generator and discriminator of the CGAN network by incorporating both the multi-scale feature fusion module and attention mechanism module. We tested our method on a dataset containing both infrared and near-infrared images, which retains more detailed features while also preserving the advantages of existing infrared images. The experimental results show that our proposed method achieved a peak signal-to-noise ratio (PSNR) of 16.5342 dB and a structural similarity index (SSIM) of 0.6385 on an RGB-NIR (Red, Green, Blue-Near Infrared) testing dataset, representing a 5% and 13% improvement over the original CGAN network, respectively. These results demonstrate the effectiveness of our proposed algorithm in addressing the issues of color leakage and unclear semantics in the original network. The proposed method in this paper is not only applicable to infrared image colorization but can also be widely applied to the colorization of remote sensing and CT images.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
eISSN:  
2076-3417  
入藏号:  
WOS:000977484100001  
Indexed Date:  
2023-05-21  
Institution Address:  
Univ Sci & Technol Beijing, Natl Ctr Mat Serv Safety, Beijing 100083, Peoples R China  
Institution Address:  
Southern Marine Sci & Engn Guangdong Lab Zhuhai, Zhuhai 519082, Peoples R China  
Institution Address:  
HBIS Mat Inst, 385 South Sports St, Shijiazhuang 050023, Peoples R China  
Institution Address:  
Hesteel Grp Tangsteel Co, 9 Binhe Rd, Tangshan 063000, Peoples R China  
Institution:  
Institution:  
Institution:  
HBIS Mat Inst  
Institution:  
Hesteel Grp Tangsteel Co

### 183. SR-DSFF and FENet-ReID: A Two-Stage Approach for Cross Resolution Person

摘要: In real-life scenarios, the accuracy of person re-identification (Re-ID) is subject to the limitation of camera hardware conditions and the change of image resolution caused by factors such as camera focusing errors. People call this problem cross-resolution person Re-ID. In this paper, we improve the recognition accuracy of cross-resolution person Re-ID by enhancing the image enhancement network and feature extraction network. Specifically, we treat cross-resolution person Re-ID as a two-stage task: the first stage is the image enhancement stage, and we propose a Super-Resolution Dual-Stream Feature Fusion sub-network, named SR-DSFF, which contains SR module and DSFF module. The SR-DSFF utilizes the SR module recovers the resolution of the low-resolution (LR) images and then obtains the feature maps of the LR images and super-resolution (SR) images, respectively, through the dual-stream feature fusion with learned weights extracts and fuses feature maps from LR and SR images in the DSFF module. At the end of SR-DSFF, we set a transposed convolution to visualize the feature maps into images. The second stage is the feature acquisition stage. We design a global-local feature extraction network guided by human pose estimation, named FENet-ReID. The FENet-ReID obtains the final features through multistage feature extraction and multiscale feature fusion for the Re-ID task. The two stages complement each other, making the final pedestrian feature representation has the advantage of accurate identification compared with other methods. Experimental results show that our method improves significantly compared with some state-of-the-art methods.  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
3  
ISSN:  
1687-5265  
eISSN:  
1687-5273  
入藏号:  
WOS:000831011600012  
PubMed ID:  
35837221  
Indexed Date:  
2022-08-04  
Institution Address:  
Jiangxi Univ Sci & Technol, Ganzhou 341000, Jiangxi, Peoples R China  
Institution Address:  
Xian Zhongtie Rail Transit Co Ltd, Xian 710000, Shaanxi, Peoples R China  
Institution:  
Institution:  
Xian Zhongtie Rail Transit Co Ltd

### 184. GEA-MSNet: A Novel Model for Segmenting Remote Sensing Images of Lakes

摘要: The decrease in lake area has garnered significant attention within the global ecological community, prompting extensive research in remote sensing and computer vision to accurately segment lake areas from satellite images. However, existing image segmentation models suffer from poor generalization performance, the imprecise depiction of water body edges, and the inadequate inclusion of water body segmentation information. To address these limitations and improve the accuracy of water body segmentation in remote sensing images, we propose a novel GEA-MSNet segmentation model. Our model incorporates a global efficient attention module (GEA) and multi-scale feature fusion to enhance the precision of water body delineation. By emphasizing global semantic information, our GEA-MSNet effectively learns image features from remote sensing data, enabling the accurate detection and segmentation of water bodies. This study makes three key contributions: firstly, we introduce the GEA module within the encode framework to aggregate shallow feature semantics for the improved classification accuracy of lake pixels; secondly, we employ a multi-scale feature fusion structure during decoding to expand the acceptance domain for feature extraction while prioritizing water body features in images; thirdly, extensive experiments are conducted on both scene classification datasets and Tibetan Plateau lake datasets with ablation experiments validating the effectiveness of our proposed GEA module and multi-scale feature fusion structure. Ultimately, our GEA-MSNet model demonstrates exceptional performance across multiple datasets with an average intersection ratio union (mIoU) improved to 75.49%, recall enhanced to 83.79%, pixel accuracy (PA) reaching 90.21%, and the f1-score significantly elevated to 83.25%.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001182599500001  
Indexed Date:  
2024-03-30  
Institution Address:  
Inner Mongolia Univ, Sch Elect Informat Engn, Hohhot 010021, Peoples R China  
Institution Address:  
Minist Educ China, Collaborat Innovat Ctr Grassland Ecol Secur, Hohhot 010021, Peoples R China  
Institution Address:  
Inner Mongolia Univ, Sch Ecol & Environm, Hohhot 010021, Peoples R China  
Institution:  
Institution:  
Minist Educ China  
Institution:

### 185. LSKANet: Long Strip Kernel Attention Network for Robotic Surgical Scene

摘要: Surgical scene segmentation is a critical task in Robotic-assisted surgery. However, the complexity of the surgical scene, which mainly includes local feature similarity (e.g., between different anatomical tissues), intraoperative complex artifacts, and indistinguishable boundaries, poses significant challenges to accurate segmentation. To tackle these problems, we propose the Long Strip Kernel Attention network (LSKANet), including two well-designed modules named Dual-block Large Kernel Attention module (DLKA) and Multiscale Affinity Feature Fusion module (MAFF), which can implement precise segmentation of surgical images. Specifically, by introducing strip convolutions with different topologies (cascaded and parallel) in two blocks and a large kernel design, DLKA can make full use of region- and strip-like surgical features and extract both visual and structural information to reduce the false segmentation caused by local feature similarity. In MAFF, affinity matrices calculated from multiscale feature maps are applied as feature fusion weights, which helps to address the interference of artifacts by suppressing the activations of irrelevant regions. Besides, the hybrid loss with Boundary Guided Head (BGH) is proposed to help the network segment indistinguishable boundaries effectively. We evaluate the proposed LSKANet on three datasets with different surgical scenes. The experimental results show that our method achieves new state-of-the-art results on all three datasets with improvements of 2.6%, 1.4%, and 3.4% mIoU, respectively. Furthermore, our method is compatible with different backbones and can significantly increase their segmentation accuracy. Code is available at https://github.com/YubinHan73/LSKANet.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0278-0062  
eISSN:  
1558-254X  
入藏号:  
WOS:001196733400006  
PubMed ID:  
38015689  
Indexed Date:  
2024-07-02  
Institution Address:  
Hunan Univ, Coll Elect & Informat Engn, Natl Engn Res Ctr Robot Visual Percept & Control T, Changsha 410082, Hunan, Peoples R China  
Institution Address:  
Hunan Univ, Int Sci & Technol Innovat Cooperat Base Biomed Ima, Changsha 410082, Hunan, Peoples R China  
Institution Address:  
Univ New South Wales, Sch Comp Sci & Engn, Sydney, NSW 2052, Australia  
Institution:  
Institution:  
Institution:

### 186. Dynamic adaptive scaling network for camouflaged object detection

摘要: Recently, accurately segmenting objects from backgrounds has become a significant challenge for Camouflaged Object Detection. Currently, traditional single-view approaches have limitations in predicting the boundary of camouflaged objects. Moreover, a fixed scaling ratio fails to fully explore imperceptible clues between camouflaged objects and background surroundings. To overcome these obstacles, we propose a triplet network via dynamic adaptive scaling strategy, named Dynamic Adaptive Scaling Network for Camouflaged Object Detection(DASNet). It mimics the behavior of humans dynamically adjusting observation distances when observing indistinct objects to zoom in and out images. Specifically, we design a dynamic adaptive scaling module to obtain the optimal scaling ratios for each view. In addition, we design the multi-view feature fusion module and the multi-scale feature fusion module to guide the learning of multi-view and multi-scale features and enhance feature representation. Extensive experiments demonstrate that the proposed DASNet outperforms 24 other representative methods on three public datasets.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001385222600001  
Indexed Date:  
2025-01-03  
Institution Address:  
Anhui Jianzhu Univ, Sch Elect & Informat Engn, Hefei 230601, Anhui, Peoples R China  
Institution:

### 187. An Ultra-Lightweight and Ultra-Fast Abnormal Target Identification

摘要: The faults caused by abnormal transmission line targets have become the main reason for the outage of transmission lines. This has brought great economic losses to society. At present, few abnormal target networks can be detected in real-time due to the restrictions of the bulky traditional model and the computing power of hardware. In order to accurately and quickly identify abnormal transmission line targets, we propose an ultra-lightweight and ultra-fast target identification network based on adaptive feature fusion in this paper, which adopts a hardware-friendly ultra-lightweight and ultra-fast backbone network to realize feature extraction. Combined with an adaptive feature fusion path aggregation network, multi-scale feature fusion is realized efficiently. The adaptive feature fusion Path Aggregation Network method includes pixel-level adaptive feature fusion and channel-level adaptive feature fusion, which can learn pixel correlation and channel importance independently. It is also beneficial to complex shape detection in complex background of a transmission line scene. In order to solve the problem of imbalance between positive and negative samples in the dataset of abnormal transmission line targets, we adopt a new loss function considering the confidence of positive and negative samples. The experiment results show that the model size is only 3.6MB, the precision (mAP) of identification for abnormal transmission line targets (4 categories) can be 61.5%, and the running time on intelligent edge identification devices is only 5 ms, with excellent performance for each indicator.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
12  
在中国科学引文数据库中的被引频次:  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
16  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:000709128900121  
Indexed Date:  
2021-11-11  
Institution Address:  
Hunan Univ, Coll Elect & Informat Engn, Changsha 410082, Peoples R China  
Institution Address:  
China Elect Power Res Inst Co Ltd, Beijing 100192, Peoples R China  
Institution Address:  
North China Elect Power Univ, Sch Elect & Elect Engn, Beijing 102206, Peoples R China  
Institution:  
Institution:  
China Elect Power Res Inst Co Ltd  
Institution:

### 188. Unveiling camouflaged and partially occluded colorectal polyps:

摘要: Background: Segmenting colorectal polyps presents a significant challenge due to the diverse variations in their size, shape, texture, and intricate backgrounds. Particularly demanding are the so-called "camouflaged"polyps, which are partially concealed by surrounding tissues or fluids, adding complexity to their detection. Methods: We present CPSNet, an innovative model designed for camouflaged polyp segmentation. CPSNet incorporates three key modules: the Deep Multi -Scale -Feature Fusion Module, the Camouflaged Object Detection Module, and the Multi -Scale Feature Enhancement Module. These modules work collaboratively to improve the segmentation process, enhancing both robustness and accuracy. Results: Our experiments confirm the effectiveness of CPSNet. When compared to state-of-the-art methods in colon polyp segmentation, CPSNet consistently outperforms the competition. Particularly noteworthy is its performance on the ETIS-LaribPolypDB dataset, where CPSNet achieved a remarkable 2.3% increase in the Dice coefficient compared to the Polyp -PVT model. Conclusion: In summary, CPSNet marks a significant advancement in the field of colorectal polyp segmentation. Its innovative approach, encompassing multi -scale feature fusion, camouflaged object detection, and feature enhancement, holds considerable promise for clinical applications.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:001202557900001  
PubMed ID:  
38394804  
Indexed Date:  
2024-04-25  
Institution Address:  
North China Univ Technol, Sch Informat Technol, Beijing 100041, Peoples R China  
Institution Address:  
Sun Yat sen Univ, Sch Intelligent Syst Engn, Guangzhou 510335, Peoples R China  
Institution Address:  
Second Peoples Hosp, Dept Gastroenterol, Changzhi 046000, Shanxi, Peoples R China  
Institution Address:  
Southern Med Univ, Sch Biomed Engn, Guangzhou 510335, Peoples R China  
Institution Address:  
Peoples Liberat Army Gen Hosp, Dept Emergency, Beijing 100853, Peoples R China  
Institution:  
Institution:  
Institution:  
Second Peoples Hosp  
Institution:  
Institution:

### 189. TL-YOLO: Foreign-Object Detection on Power Transmission Line Based on

摘要: Foreign objects on power transmission lines carry a significant risk of triggering large-scale power interruptions which may have serious consequences for daily life if they are not detected and handled in time. To accurately detect foreign objects on power transmission lines, this paper proposes a TL-Yolo method based on the Yolov8 framework. Firstly, we design a full-dimensional dynamic convolution (ODConv) module as a backbone network to enhance the feature extraction capability, thus retaining richer semantic content and important visual features. Secondly, we present a feature fusion framework combining a weighted bidirectional feature pyramid network (BiFPN) and multiscale attention (MSA) module to mitigate the degradation effect of multiscale feature representation in the fusion process, and efficiently capture the high-level feature information and the core visual elements. Thirdly, we utilize a lightweight GSConv cross-stage partial network (GSCSP) to facilitate efficient cross-level feature fusion, significantly reducing the complexity and computation of the model. Finally, we employ the adaptive training sample selection (ATSS) strategy to balance the positive and negative samples, and dynamically adjust the selection process of the training samples according to the current state and performance of the model, thus effectively reducing the object misdetection and omission. The experimental results show that the average detection accuracy of the TL-Yolo method reaches 91.30%, which is 4.20% higher than that of the Yolov8 method. Meanwhile, the precision and recall metrics of our method are 4.64% and 3.53% higher than those of Yolov8. The visualization results also show the superior detection performance of the TL-Yolo algorithm in real scenes. Compared with the state-of-the-art methods, our method achieves higher accuracy and speed in the detection of foreign objects on power transmission lines.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
eISSN:  
2079-9292  
入藏号:  
WOS:001210707000001  
Indexed Date:  
2024-05-10  
Institution Address:  
Nantong Univ, Sch Transportat & Civil Engn, Nantong 226019, Peoples R China  
Institution Address:  
Nantong Univ, Sch Elect Engn, Nantong 226004, Peoples R China  
Institution:  
Institution:

### 190. Detection of Dense Citrus Fruits by Combining Coordinated Attention and

摘要: The accuracy detection of individual citrus fruits in a citrus orchard environments is one of the key steps in realizing precision agriculture applications such as yield estimation, fruit thinning, and mechanical harvesting. This study proposes an improved object detection YOLOv5 model to achieve accurate the identification and counting of citrus fruits in an orchard environment. First, the latest visual attention mechanism coordinated attention module (CA) was inserted into an improved backbone network to focus on fruit-dense regions to recognize small target fruits. Second, an efficient two-way cross-scale connection and weighted feature fusion BiFPN in the neck network were used to replace the PANet multiscale feature fusion network, giving effective feature corresponding weights to fully fuse the high-level and bottom-level features. Finally, the varifocal loss function was used to calculate the model loss for better model training results. The results of the experiments on four varieties of citrus trees showed that our improved model proposed to this study could effectively identify dense small citrus fruits. Specifically, the recognized AP (average precision) reached 98.4%, and the average recognition time was 0.019 s per image. Compared with the original YOLOv5 (including deferent variants of n, s, m, l, and x), the increase in the average accuracy precision of the improved YOLOv5 ranged from 7.5% to 0.8% while maintaining similar average inference time. Four different citrus varieties were also tested to evaluate the generalization performance of the improved model. The method can be further used as a part in a vision system to provide technical support for the real-time and accurate detection of multiple fruit targets during mechanical picking in citrus orchards.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
16  
在 BIOSIS Citation Index 中的被引频次:  
5  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
18  
eISSN:  
2076-3417  
入藏号:  
WOS:000823669800001  
Indexed Date:  
2022-07-20  
Institution Address:  
Guangxi Normal Univ, Guangxi Key Lab Multisource Informat Min & Secur, Coll Comp Sci & Engn, Guilin 541004, Peoples R China  
Institution Address:  
Guangxi Acad Specialty Crops, Guangxi Citrus Breeding & Cultivat Engn Technol C, Guilin 541004, Peoples R China  
Institution:  
Institution:  
Guangxi Acad Specialty Crops

### 191. Multiple Defect Detection of Insulators Based on Improved YOLOv5

摘要: Aiming at such issues as low detection accuracy and slow speed of multi-defect detection of insulator,a kind of improved YOLOv5 accurate detection algorithm for multi-defect detection of insulator(YOLOv5-GSEM)is proposed. Firstly,GhostNet structure is introduced to replace C3 module of the original YOLOv5 backbone network to improve the network computing speed. After SPPF,the non-parametric attention module SimAM is introduced to enhance the effective features and suppress the interference features. Secondly,the enhanced feature pyramid network (EFPN)and the multiscale feature fusion network(MFFN)are introduced to fully integrate multiscale features and improve the detection accuracy of multi-defects of insulator by the network. The experimental results show that the average of the accuracy of the model(mAP0.5)proposed in this paper reaches 87.8%,which is 2.7% higher than that of YOLOv5 algorithm,and the detection speed is increased by 4.6%. The proposed network provides a more effective method for the detection of various defects of insulator.  
摘要:  
针对绝缘子多缺陷检测精度低、检测速度慢的问题,提出一种改进YOLOv5准确判别绝缘子多缺陷检测算法(YOLOv5-GSEM)。首先通过引入GhostNet结构替换原始网络YOLOv5主干网络C3模块,提升网络运算速度;并在SPPF后引入无参注意力模块SimAM,增强有效特征,抑制干扰特征;其次引入增强特征金字塔网络(EFPN)和多尺度特征融合网络(multiscale feature fusion network,MFFN),充分融合多尺度特征,提升网络对绝缘子多缺陷的检测精度。实验结果表明,文中提出的模型平均精度均值(mAP0.5)达到87.8%,较YOLOv5算法提升了2.7%,检测速度提升了4.6%,该网络的提出为绝缘子多种缺陷检测问题提供一种更有效的方法。  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1001-1609  
入藏号:  
CSCD:7860707  
Indexed Date:  
2025-02-15  
Institution Address:  
三峡大学电气与新能源学院, 宜昌, 湖北 443002, 中国  
Institution Address:  
三峡大学电气与新能源学院;;三峡大学, ;;湖北省输电线路工程技术研究中心, 宜昌;;宜昌, 湖北;;湖北 443002;;443002, 中国  
Institution Address:  
三峡大学计算机与信息学院, 宜昌, 湖北 443002, 中国  
Institution Address:  
College of Electrical Engineering and New Energy,China Three Gorges University, Yichang, Hubei 443002, China  
Institution Address:  
College of Electrical Engineering and New Energy,China Three Gorges University;;Three Gorges University, ;;Hubei Provincial Engineering Technology Research Center for Power Transmission Line, Yichang;;Yichang, Hubei;;Hubei 443002;;443002  
Institution Address:  
College of Computer and Information,China Three Gorges University, Yichang, Hubei 443002, China  
Institution:  
三峡大学电气与新能源学院  
Institution:  
三峡大学电气与新能源学院;;三峡大学  
Institution:  
三峡大学计算机与信息学院  
Institution:  
College of Electrical Engineering and New Energy,China Three Gorges University  
Institution:  
College of Electrical Engineering and New Energy,China Three Gorges University;;Three Gorges University  
Institution:  
College of Computer and Information,China Three Gorges University

### 192. Edge detection for 3D point clouds via locally max-angular gaps

摘要: Numerous point-cloud-based applications, including surface reconstruction and completion, greatly benefit from the utilization of edge features, which play a crucial role in structuring the target shape. Nevertheless, the current limitations of point cloud edge detection techniques in effectiveness and efficiency have motivated us to develop a more robust 3D edge detection method. In this paper, we introduce a simple yet effective descriptor for evaluating the edge level of 3D point clouds. This descriptor is defined as the variance of the max-angular gaps between the target point and its k nearest neighboring points. Leveraging this descriptor, we further propose a multiscale feature fusion strategy that well detect the edge features for the target point cloud. Experimental results demonstrate the superiority of our method over other comparison methods in terms of effectiveness and stability in preserving point cloud edges. Moreover, we investigate the applications of our method in point cloud simplification and registration to provide additional insights into its potential use in some point-cloud-based downstream tasks.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001105596100001  
Indexed Date:  
2023-12-17  
Institution Address:  
Sun Yat Sen Univ, Sch Aeronaut & Astronaut, Shenzhen 518170, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Sch Elect & Commun Engn, Shenzhen 518170, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Artificial Intelligence, Nanjing 210044, Peoples R China  
Institution:  
Institution:  
Institution:

### 193. Dynamic balanced multi-scale feature fusion for colorectal polyp

摘要: Colorectal cancer is one of the most prevalent diseases,and accurate colorectal polyp segmentation can aid physicians in early prevention.However,in the process of segmentation,colorectal polyp images present several challenges,such as low contrast,varied shapes of lesions,and randomized location.Moreover,with large parameters,the Unet network does not yield high segmentation accuracy.Therefore,an improved Unet algorithm based on dynamic balanced multiscale feature fusion was proposed.This algorithm took Unet as the main body and combined the atrous spatial pyramid pooling module (ASPP).A channel shuffle inception (CSI) module and a group inception(GI) module were also put forth to improve the convolution block of the codec,reduce the amount of network parameters,and improve the model's characterization ability.Additionally,a residual pyramid split attention module(RPSA) was presented for the skip connection of the codec,balancing the channel information in the skip connection,and improving the overall network split performance.Experimental results showed that this method could not only outperform other methods in terms of segmentation effect,but also significantly reduce the number of parameters,thereby proving its effectiveness.  
摘要:  
结直肠癌作为最常见的疾病之一,精准的结直肠息肉分割可辅助医师对其进行早期预防。然而,在分割过程中,结直肠息肉图像存在对比度较低、病灶形状不一、位置随机化等问题,而且Unet网络参数量较大但分割精度不高。因此,提出了一种基于动态平衡多尺度特征融合的Unet改进算法,以Unet为主体,结合空洞空间卷积池化金字塔模块(ASPP)提高Unet深层次特征的多样性;提出通道打乱多尺度特征融合模块(CSI)和分组多尺度特征融合模块(GI)对编解码器的卷积块进行改进,降低整体网络参数量同时提高模型的表征能力,并提出残差金字塔拆分注意力模块(RPSA)用于编解码器的跳跃连接,平衡跳跃连接中的通道信息,提高整体网络的分割性能。实验结果表明,该方法不仅在分割效果上优于其他方法,还大幅减少了参数量,证明了其有效性。  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
1  
ISSN:  
2095-302X  
入藏号:  
CSCD:7462060  
Indexed Date:  
2023-08-12  
Institution Address:  
桂林理工大学信息科学与工程学院;;广西嵌入式技术与智能系统重点实验室, ;;广西嵌入式技术与智能系统重点实验室, 桂林;;桂林, 广西;;广西 541004;;541004, 中国  
Institution Address:  
桂林理工大学信息科学与工程学院, 桂林, 广西 541004, 中国  
Institution Address:  
School of Information Science and Engineering,Guilin University of Technology;;Guangxi Key Laboratory of Embedded Technology and Intelligent System, ;;Guangxi Key Laboratory of Embedded Technology and Intelligent System, Guilin;;Guilin, Guangxi;;Guangxi 541004;;541004  
Institution Address:  
School of Information Science and Engineering,Guilin University of Technology, Guilin, Guangxi 541004, China  
Institution:  
桂林理工大学信息科学与工程学院;;广西嵌入式技术与智能系统重点实验室  
Institution:  
桂林理工大学信息科学与工程学院  
Institution:  
School of Information Science and Engineering,Guilin University of Technology;;Guangxi Key Laboratory of Embedded Technology and Intelligent System  
Institution:  
School of Information Science and Engineering,Guilin University of Technology

### 194. LCDNet: Lightweight Change Detection Network With Dual-Attention

摘要: Deep-learning (DL)-based change detection (CD) techniques have recently become increasingly complex to produce more accurate detection results. However, the increase in complexity leads to reduced efficiency and limits the application of DL-based CD techniques in domains that require real-time performance. To this end, a lightweight CD network (LCDNet) is proposed to accurately recognize changes in remote-sensing (RS) image pairs while maintaining high efficiency. First, a focus module is utilized at the beginning of the encoding layer for the downsampling operation, which reduces the computation of the model and the loss of information. Then, a depthwise (DW) convolution-based efficient extraction block (EEB) is designed by stacking different sizes of convolution kernels for the effective extraction of change features under different receptive fields. Next, a dual-attention guidance module (DAGM) is designed to guide the encoder in processing and selectively aggregating information related to changes. Lastly, a multiscale feature fusion module (MFFM) with low parameters is proposed that combines feature maps of different scales to exploit their complementary information. Compared with other state-of-the-art (SOTA) methods, the proposed LCDNet only requires approximately 0.83 M Params, 2.03 G FLOPs, and 3.03 ms inference time (It) to remarkably surpass them in terms of accuracy. Moreover, compared with other dual-attention and multiscale fusion modules, the proposed DAGM and MFFM are more effective and efficient. The source code will be made available at https://github.com/sjl2023/LCDNet.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001136775600058  
Indexed Date:  
2024-02-08  
Institution Address:  
Henan Univ, Sch Artificial Intelligence, Zhengzhou 450046, Peoples R China  
Institution Address:  
Weinan Normal Univ, Sch Phys & Elect Engn, Weinan 714099, Peoples R China  
Institution:  
Institution:

### 195. Multiscale Traffic Sign Detection Method in Complex Environment Based on

摘要: Traffic sign detection is a challenging problem in the field of unmanned driving, particularly important in complex environments. We propose a method, based on the improved You only look once (YOLO) v4, to detect and recognize multiscale traffic signs in complex environments. This method employs an image preprocessing module that can classify and denoize images of complex environments and then input the images into the improved YOLOv4. We also design an improved feature pyramid structure to replace the original feature pyramid of YOLOv4. This structure uses an adaptive feature fusion module and a multiscale feature transfer mechanism to reduce putative information loss in the feature map generation process and improve the information transfer between deep and shallow features, enhancing the representation ability of feature pyramids. Finally, we use EIOU LOSS and Cluster-NMS to further improve the model performance. The experimental results on the fusion of Tsinghua-Tencent 100 K and our collected dataset show that the proposed method achieves an mAP of 81.78%. Compared to existing methods, our method demonstrates its superiority with regard to traffic sign detection.  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
1687-5265  
eISSN:  
1687-5273  
入藏号:  
WOS:000877869100003  
PubMed ID:  
36317077  
Indexed Date:  
2022-11-13  
Institution Address:  
Northeast Forestry Univ, Coll Informat & Comp Engn, Harbin 150040, Peoples R China  
Institution:

### 196. Infrared Small UAV Target Detection Based on Depthwise Separable

摘要: Unmanned aerial vehicles (UAVs) have been widely applied in military and civilian fields, but they also pose great threats to restricted areas, such as densely populated areas and airports. Thermal infrared (IR) imaging technology is capable of monitoring UAVs at a long range in both day and night conditions. Therefore, the anti-UAV technology based on thermal IR imaging has attracted growing attention. However, the images acquired by IR sensors often suffer from small and dim targets, as well as heavy background clutter and noise. Conventional detection methods usually have a high false alarm rate and low detection accuracy. This article proposes a detection method that formulates the UAV detection as predicting the residual image (i.e., background, clutter, and noise) by learning the nonlinear mapping from the input image to the residual image. The UAV target image is obtained by subtracting the residual image from the input IR image. The constructed end-to-end U-shaped network exploits the depthwise separable residual dense blocks in the encoder stage to extract the abundant hierarchical features. Besides, the multiscale feature fusion and representation block is introduced to fully aggregate multiscale features from the encoder layers and intermediate connection layers at the same scale, as well as the decoder layers at different scales, to better reconstruct the residual image in the decoder stage. In addition, the global residual connection is adopted in the proposed network to provide long-distance information compensation and promote gradient backpropagation, which further improves the performance in reconstructing the image. The experimental results show that the proposed method achieves favorable detection performance in real-world IR images and outperforms other state-of-the-art methods in terms of quantitative and qualitative evaluation metrics.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
4  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
48  
被引频次合计:  
52  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000845062100008  
Indexed Date:  
2022-09-03  
Institution Address:  
Xidian Univ, Sch Comp Sci & Technol, Xian 710071, Peoples R China  
Institution Address:  
Huazhong Univ Sci & Technol, Sch Artificial Intelligence & Automat, Wuhan 430074, Peoples R China  
Institution Address:  
Zhoukou Normal Univ, Coll Phys & Telecommun Engn, Zhoukou 466001, Peoples R China  
Institution:  
Institution:  
Institution:

### 197. MRFA-Net: Kidney Segmentation Method Based on Multi-Scale Feature Fusion

摘要: For the characterization of the kidney segmentation task, this paper proposes a self-supervised kidney segmentation method based on multi-scale feature fusion and residual full attention, named MRFA-Net. In this study, we introduce the multi-scale feature fusion module to extract multi-scale information of kidneys from abdominal CT slices; additionally, the residual full-attention convolution module is designed to handle the multi-scale information of kidneys by introducing a full-attention mechanism, thus improving the segmentation results of kidneys. The Dice coefficient on the Kits19 dataset reaches 0.972. The experimental results demonstrate that the proposed method achieves good segmentation performance compared to other algorithms, effectively enhancing the accuracy of kidney segmentation.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001191728100001  
Indexed Date:  
2024-04-03  
Institution Address:  
Kunming Univ Sci & Technol, Fac Informat Engn & Automat, Kunming 650500, Peoples R China  
Institution Address:  
Kunming Univ Sci & Technol, Fac Modern Agr Engn, Kunming 650500, Peoples R China  
Institution Address:  
Yunnan Prov Key Lab Comp, Kunming 650500, Peoples R China  
Institution:  
Institution:  
Institution:  
Yunnan Prov Key Lab Comp

### 198. Person re-identification based on multi-scale feature fusion and

摘要: Person re-identification is an image retrieval technique for person in real scenes. Due to factors such as camera angle, lighting, and occlusion, there is a high intra-class variation in the representation of a specific sample. Furthermore, discriminative local regions such as hats and shoes are often ignored, resulting in some useful local information being unable to be used for retrieval. In this paper, a multi-scale feature fusion network model combining global and local features is proposed. The network is built with four stacked building block, where multi-scale features are assigned with different weights and fused according to the output conditions of each branch. In addition, a multi-attention mechanism network is combined with the multi-scale feature fusion in this paper. This method aims to enable the network to model the relation between input images, so as to effectively aggregate the features of neighbour person samples to obtain a more robust image representation. Experimental results show that the retrieval performance can be improved by the proposed method.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001156916200053  
Indexed Date:  
2024-02-27  
Institution Address:  
Soochow Univ, Sch Elect & Informat Engn, Suzhou 215006, Peoples R China  
Institution:

### 199. Muti-Frame Point Cloud Feature Fusion Based on Attention Mechanisms for

摘要: Continuous frames of point-cloud-based object detection is a new research direction. Currently, most research studies fuse multi-frame point clouds using concatenation-based methods. The method aligns different frames by using information on GPS, IMU, etc. However, this fusion method can only align static objects and not moving objects. In this paper, we proposed a non-local-based multi-scale feature fusion method, which can handle both moving and static objects without GPS- and IMU-based registrations. Considering that non-local methods are resource-consuming, we proposed a novel simplified non-local block based on the sparsity of the point cloud. By filtering out empty units, memory consumption decreased by 99.93%. In addition, triple attention is adopted to enhance the key information on the object and suppresses background noise, further benefiting non-local-based feature fusion methods. Finally, we verify the method based on PointPillars and CenterPoint. Experimental results show that the mAP of the proposed method improved by 3.9% and 4.1% in mAP compared with concatenation-based fusion modules, PointPillars-2 and CenterPoint-2, respectively. In addition, the proposed network outperforms powerful 3D-VID by 1.2% in mAP.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
10  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
eISSN:  
1424-8220  
入藏号:  
WOS:000867320500001  
PubMed ID:  
36236572  
Indexed Date:  
2022-10-21  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100190, Peoples R China  
Institution Address:  
Chinese Acad Sci, Key Lab Technol Geospatial Informat Proc & Applic, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 101408, Peoples R China  
Institution:  
Institution:  
Institution:

### 200. YOLOv7F: Enhanced YOLOv7 With Guided Feature Fusion

摘要: A general object detector has a high misdetection rate for small objects. Although many small-object detectors consider the insufficient representation of objects, their performance in detecting very tiny objects with a strong similarity to other objects and backgrounds in aerial images remains poor. In this study, we analyze the misalignments of spatial and semantic information of features due to resizing, involving interpolation and pooling operations conducted before multi-scale feature fusion. Additionally, as a learning target, the objectness loss uses IoU values, which are sensitive to the minute distance differences between predicted small objects in the detector and the ground-truth data. Therefore, the neck and head architecture of the proposed You Only Look Once version 7 Fusion (YOLOv7F) model is redesigned to be suitable for small-object detection. The YOLOv7F model includes the Deformable Feature Fusion (DFF) module, which aligns the features based on the guided features, and the Objectness Refinement Head (ORH) model, which refines the predicted objectness score. The YOLOv7F model achieved 63.9% $mAP\_{0.5}$ performance and led to a 4.1% improvement compared to the YOLOv7X model on the AI-TODv2, where small objects account for 98.1% of the all instances. In the VisDrone2019-DET dataset, where 32.0% of instances are larger than a medium-sized object, YOLOv7F model achieved an $mAP\_{0.5}$ of 63.9%, a 2.0% improvement compared to the YOLOv7X model.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001362127900026  
Indexed Date:  
2024-12-11  
Institution Address:  
Hanwha Syst Co Ltd, Seongnam 13524, Gyeonggi, South Korea  
Institution:

### 201. 3D Object Detection Based on Attention and Multi-Scale Feature Fusion

摘要: Three-dimensional object detection in the point cloud can provide more accurate object data for autonomous driving. In this paper, we propose a method named MA-MFFC that uses an attention mechanism and a multi-scale feature fusion network with ConvNeXt module to improve the accuracy of object detection. The multi-attention (MA) module contains point-channel attention and voxel attention, which are used in voxelization and 3D backbone. By considering the point-wise and channel-wise, the attention mechanism enhances the information of key points in voxels, suppresses background point clouds in voxelization, and improves the robustness of the network. The voxel attention module is used in the 3D backbone to obtain more robust and discriminative voxel features. The MFFC module contains the multi-scale feature fusion network and the ConvNeXt module; the multi-scale feature fusion network can extract rich feature information and improve the detection accuracy, and the convolutional layer is replaced with the ConvNeXt module to enhance the feature extraction capability of the network. The experimental results show that the average accuracy is 64.60% for pedestrians and 80.92% for cyclists on the KITTI dataset, which is 1.33% and 2.1% higher, respectively, compared with the baseline network, enabling more accurate detection and localization of more difficult objects.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
14  
eISSN:  
1424-8220  
入藏号:  
WOS:000803269200001  
PubMed ID:  
35632344  
Indexed Date:  
2022-06-09  
Institution Address:  
Xinjiang Univ, Sch Informat Sci & Engn, Urumqi 830046, Peoples R China  
Institution:

### 202. Multiple myeloma segmentation net (MMNet): an encoder-decoder-based deep

摘要: Background: Patients with multiple myeloma (MM), a malignant disease involving bone marrow plasma cells, shows significant susceptibility to bone degradation, impairing normal hematopoietic function. The accurate and effective segmentation of MM lesion areas is crucial for the early detection and diagnosis of myeloma. However, the presence of complex shape variations, boundary ambiguities, and multiscale lesion areas, ranging from punctate lesions to extensive bone damage, presents a formidable challenge in achieving precise segmentation. This study thus aimed to develop a more accurate and robust segmentation method for MM lesions by extracting rich multiscale features. Methods: In this paper, we propose a novel, multiscale feature fusion encoding-decoding model architecture specifically designed for MM segmentation. In the encoding stage, our proposed multiscale feature extraction module, dilated dense connected net (DCNet), is employed to systematically extract multiscale features, thereby augmenting the model's sensing field. In the decoding stage, we propose the CBAM-atrous spatial pyramid pooling (CASPP) module to enhance the extraction of multiscale features, enabling the model to dynamically prioritize both channel and spatial information. Subsequently, these features are concatenated with the final output feature map to optimize segmentation outcomes. At the feature fusion bottleneck layer, we incorporate the dynamic feature fusion (DyCat) module into the skip connection to dynamically adjust feature extraction parameters and fusion processes. Results: We assessed the efficacy of our approach using a proprietary dataset of MM, yielding notable advancements. Our dataset comprised 753 magnetic resonance imaging (MRI) two-dimensional (2D) slice images of the spinal regions from 45 patients with MM, along with their corresponding ground truth labels. These images were primarily obtained from three sequences: T1-weighted imaging (T1WI), T2-weighted imaging (T2WI), and short tau inversion recovery (STIR). Using image augmentation techniques, we expanded the dataset to 3,000 images, which were employed for both model training and prediction. Among these, 2,400 images were allocated for training purposes, while 600 images were reserved for validation and testing. Our method showed increase in the intersection over union (IoU) and Dice coefficients by 7.9 and 6.7 percentage points, respectively, as compared to the baseline model. Furthermore, we performed comparisons with alternative image segmentation methodologies, which confirmed the sophistication and efficacy of our multiscale features from images and enhance the correlation between channel and spatial information. Furthermore, a systematic evaluation of the proposed network architecture was conducted on a self- constructed, limited dataset. This endeavor holds promise for offering valuable insights into the development of algorithms for future clinical applications.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2223-4292  
eISSN:  
2223-4306  
入藏号:  
WOS:001354315400011  
PubMed ID:  
39429589  
Indexed Date:  
2024-11-30  
Institution Address:  
Chongqing Jiaotong Univ, Sch Mechatron & Vehicle Engn, 66 Xuefu Ave, Chongqing 400074, Peoples R China  
Institution Address:  
Chongqing Med Univ, Dept Blood Transfus, Affiliated Hosp 1, 1 Youyi Rd, Chongqing 400074, Peoples R China  
Institution:  
Institution:

### 203. Solar cell defect detection network combining multiscale feature and

摘要: In order to improve the detection accuracy of various types of defects in the electroluminescence imaging of solar cells, a solar cell surface defect detection algorithm CMFAnet was proposed by fusing multiscale features and attention mechanism. Firstly, for the characteristics of solar cell surface defects with large scale span, an enhanced multi-scale feature fusion method was designed, whose basic unit consists of a feature alignment module and a feature fusion module connected in series, and for the feature information with different semantic levels, the feature alignment module adjusts their scales, so that these features can be fused together more easily; secondly, for the characteristics of solar cell surface defects with high level and variable geometry, a deformable ghost convolution module is designed. Secondly, for the characteristics of high degree of similarity between defective features and background features on the solar cell surface and variable geometry, a deformable ghost convolution module was designed, whose basic unit consists of feasible variant convolution, multiplexed coordinate attention mechanism, and ghost convolution; the multiplexed coordinate attention mechanism optimizes the generation of offset in the feasible variant convolution, and the introduction of ghost convolution mechanism effectively reduces the FLOPs of the network model. The experimental results show that the mAP of this paper's method reaches 91.4% on the photovoltaic cell defect anomaly detection dataset PVEL-AD, which is improved to different degrees compared to other mainstream target detection networks.  
摘要:  
为提高对太阳能电池电致发光(EL)成像各类表面缺陷的检测精度并降低漏检率,提出融合多尺度特征与注意力机制的太阳能电池表面缺陷检测算法CMFAnet。首先,针对太阳能电池表面缺陷尺度跨度大的特点,设计了增强型多尺度特征融合方法,其基本单元由特征对齐模块和特征融合模块串联组成,对于不同语义级别的特征信息,特征对齐模块通过调整它们的尺度,使这些特征更容易融合在一起;其次,针对太阳能电池表面缺陷特征与背景特征相似程度高、几何形状多变的特点,设计了可形变幽灵卷积模块,其基本单元由可形变卷积、多路坐标注意力机制和幽灵卷积(Ghost conv)组成,多路坐标注意力机制优化了可形变卷积中offset的生成,幽灵卷积机制的引入则有效降低了网络模型的计算复杂度。实验结果表明,在光伏电池缺陷异常检测数据集PVEL-AD上,本文方法的平均检测精度(mAP)达91.4%,相较其他主流目标检测网络均有不同程度的提升。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1004-924X  
入藏号:  
CSCD:7784386  
Indexed Date:  
2024-10-18  
Institution Address:  
河北工业大学人工智能与数据科学学院;;河北省控制工程技术研究中心, ;;河北省控制工程技术研究中心, ;;, 天津;;天津 300130;;300130, 中国  
Institution Address:  
河北工业大学人工智能与数据科学学院, 天津 300130, 中国  
Institution Address:  
School of Artificial Intelligence,Hebei University of Technology;;China Hebei Control Engineering Research Center, ;;China Hebei Control Engineering Research Center, ;;, Tianjin;;Tianjin 300130;;300130  
Institution Address:  
School of Artificial Intelligence,Hebei University of Technology, Tianjin 300130, China  
Institution:  
河北工业大学人工智能与数据科学学院;;河北省控制工程技术研究中心  
Institution:  
河北工业大学人工智能与数据科学学院  
Institution:  
School of Artificial Intelligence,Hebei University of Technology;;China Hebei Control Engineering Research Center  
Institution:  
School of Artificial Intelligence,Hebei University of Technology

### 204. YOLO-RMS: A Lightweight and Efficient Detector for Object Detection in

摘要: Object detection (OD) plays a critical role in interpreting optical remote sensing images (ORSIs) within the field of Earth observation. Despite significant progress has been made in natural scene OD models, applying these detectors directly to remote sensing images has not achieved the expected results due to the complexity of the scenes and the dramatic variation in shape and scale in remote sensing images. A lightweight and efficient detector based on YOLOv8n is proposed to address these challenges. First, we design Reassembly-PAN (RA-PAN) to guide the reassembly of features in multiscale feature fusion. Second, multiscale dilated attention (MSDA) is introduced after the feature fusion module to make the model focus on effective features in complex backgrounds. Finally, Shape-IoU is employed as the bounding box regression loss to make the model focus on the shape and scale of the bounding box itself and improve its localization ability. Experimental results show that the mean average precision is improved by 4.7% compared with the baseline model on the fine-grained optical remote sensing dataset SIMD, while introducing a few additional parametric quantities. In addition, the proposed method improves the mean average accuracy (mAP) on the optical remote sensing dataset DIOR and RSOD by 1.7% and 2.1%, respectively. Better detection performance is achieved compared with other mainstream models.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
3  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001283693700023  
Indexed Date:  
2024-08-14  
Institution Address:  
Huazhong Univ Sci & Technol, Sch Software Engn, Wuhan 430074, Peoples R China  
Institution:

### 205. Multi-Scale Feature Fusion with Adaptive Weighting for Diabetic

摘要: Diabetic retinopathy (DR) is the prime cause of blindness in people who suffer from diabetes. Automation of DR diagnosis could help a lot of patients avoid the risk of blindness by identifying the disease and making judgments at an early stage. The main focus of the present work is to propose a feasible scheme of DR severity level detection under the MobileNetV3 backbone network based on a multi-scale feature of the retinal fundus image and improve the classification performance of the model. Firstly, a special residual attention module RCAM for multi-scale feature extraction from different convolution layers was designed. Then, the feature fusion by an innovative operation of adaptive weighting was carried out in each layer. The corresponding weight of the convolution block is updated in the model training automatically, with further global average pooling (GAP) and division process to avoid over-fitting of the model and removing non-critical features. In addition, Focal Loss is used as a loss function due to the data imbalance of DR images. The experimental results based on Kaggle APTOS 2019 contest dataset show that our proposed method for DR severity classification achieves an accuracy of 85.32%, a kappa statistic of 77.26%, and an AUC of 0.97. The comparison results also indicate that the model obtained is superior to the existing models and presents superior classification performance on the dataset.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
17  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
17  
eISSN:  
2079-9292  
入藏号:  
WOS:000666439500001  
Indexed Date:  
2021-07-05  
Institution Address:  
Guizhou Univ, Coll Big Data & Informat Engn, Guiyang 550025, Peoples R China  
Institution:

### 206. Lightweight safety helmet wearing detection fusing coordinate attention

摘要: The existing algorithm for detecting the helmet wear by coal miners has the problem of difficulty in achieving a good balance between detection accuracy and speed. In order to solve the above problem, based on the YOLOv4 model, a lightweight model (M-YOLO) that integrates coordinate attention and multi-scale is proposed and applied in safety helmet wearing detection. This model replaces YOLOv4's feature extraction network CSPDarknet53 with a lightweight feature extraction network S-MobileNetV2 composed of a mixed coordinate attention module. It effectively improves the connection between features while reducing the number of related parameters. The model changes the parallel connection method in the original spatial pyramid pooling structure to serial connection. It effectively improves computational efficiency. The feature fusion network is improved by introducing shallow features with high-resolution and multi detail texture information. It effectively enhances the extraction of object features. Some convolutions in the original Neck structure are modified to deep separable convolutions, further reducing the model's parameter and computational complexity while ensuring detection precision. The experimental results show that compared with the YOLOv4 model, the mean average precision of the M-YOLO model is only reduced by 0.84%. But the computational complexity, parameter quantity, and model size are reduced by 74.5%, 72.8%, and 81.6%, respectively. The detection speed is improved by 53.4%. Compared to other models, the M-YOLO model achieves a good balance between accuracy and realtime performance, meeting the requirements of embedded loading and deployment on intelligent video surveillance terminals.  
摘要:  
针对现有煤矿工人安全帽佩戴检测算法存在检测精度与速度难以取得较好平衡的问题,以YOLOv4模型为基础,提出了一种融合坐标注意力与多尺度的轻量级模型M-YOLO,并将其用于安全帽佩戴检测。该模型使用融入混洗坐标注意力模块的轻量化特征提取网络S-MobileNetV2替换YOLOv4的特征提取网络CSPDarknet53,在减少相关参数量的前提下,有效改善了特征之间的联系;将原有空间金字塔池化结构中的并行连接方式改为串行连接,有效提高了计算效率;对特征融合网络进行改进,引入具有高分辨率、多细节纹理信息的浅层特征,以有效加强对检测目标特征的提取,并将原有Neck结构中的部分卷积修改为深度可分离卷积,在保证检测精度的前提下进一步降低了模型的参数量和计算量。实验结果表明,与YOLOv4模型相比,M-YOLO模型的平均精度均值仅降低了0.84%,但计算量、参数量、模型大小分别减小了74.5%,72.8%,81.6%,检测速度提高了53.4%;相较于其他模型,M-YOLO模型在准确率和实时性方面取得了良好的平衡,满足在智能视频监控终端上嵌入式加载和部署的需求。  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1671-251X  
入藏号:  
CSCD:7603070  
Indexed Date:  
2024-01-27  
Institution Address:  
内蒙古电投能源股份有限公司北露天煤矿, 霍林郭勒, 内蒙古 029200, 中国  
Institution Address:  
(徐州)中国矿业大学信息与控制工程学院, 徐州, 江苏 221116, 中国  
Institution Address:  
北京和利时数字技术有限公司, 北京 100176, 中国  
Institution Address:  
North Open-pit Coal Mine, Inner Mongolia Power Investment Energy Co., Ltd., Holingola, 029200  
Institution Address:  
School of Information and Control Engineering, China University of Mining and Technology(Xuzhou), Xuzhou, Jiangsu 221116, China  
Institution Address:  
Beijing Hollysys Digital Technology Co., Ltd., Beijing 100176, China  
Institution:  
内蒙古电投能源股份有限公司北露天煤矿  
Institution:  
(徐州)中国矿业大学信息与控制工程学院  
Institution:  
北京和利时数字技术有限公司  
Institution:  
North Open-pit Coal Mine, Inner Mongolia Power Investment Energy Co., Ltd.  
Institution:  
School of Information and Control Engineering, China University of Mining and Technology(Xuzhou)  
Institution:  
Beijing Hollysys Digital Technology Co., Ltd.

### 207. Enhanced multi-scale feature progressive network for image Deblurring

摘要: This paper tackles the problem of single image motion blur removal. Recently methods have achieved state-of-the-art results owe to multi-scale, scale-recurrent and coarse-to-fine architecture, however, the problem of image feature information extraction and information transfer between different stages has not been well solved. In this paper, first, an efficient Enhanced Multi-scale Feature Progressive Network (EMFPNet) was proposed, in order to solve the above problem, a multi-scale feature extraction module is applied in each stage to enrich the spatial features of the maps. Second, introducing a Cross-stage Feature Fusion module to solve the problem of information transmission in different stages. Third, a cross-stage attention mechanism is used to monitor and help the transmission of information. Compared to SOTA method, our method achieve 0.6% and 0.2% improvement in PSNR respectively on GoPro and HIDE datasets.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:000936302600003  
Indexed Date:  
2023-03-20  
Institution Address:  
Qingdao Univ, Coll Comp Sci & Technol, Qingdao 266071, Peoples R China  
Institution:

### 208. Face Reconstruction Based on Multiscale Feature Fusion and 3D Animation

摘要: The development of the Internet of Things and 3D technology promotes the wide application of face models in 3D animation. However, because the expression is inconsistent with the facial muscle movement, the reconstruction results may be far from the real appearance in the process of reconstructing the face appearance. Therefore, this paper proposes a character expression simulation model under the framework of 3DS Max. According to the relationship between head bones and muscles, a facial muscle motion model was established. Then, the expression simulation design of the original three-dimensional animation character "yaya" was carried out under the framework of 3DS Max technology. The experimental results of "yaya" facial expression test showed that the face simulation model using this method not only has vivid and natural expression but also conforms to the law of facial muscle movement, which provides an important reference for the construction and application of 3D face model.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1574-017X  
eISSN:  
1875-905X  
入藏号:  
WOS:000743151400002  
Indexed Date:  
2022-01-23  
Institution Address:  
Xijing Univ, Sch Design Art, Xian 710123, Shaanxi, Peoples R China  
Institution:

### 209. Mutual Learning Offline Handwritten Mathematical Expression Recognition

摘要: With complex two-dimensional structure, offline handwritten mathematical expressions is difficult to recognize due to the variable scale of their symbols and the various transformation of their writing styles. This paper proposed a mutual learning model based on multi-scale feature fusion. Firstly, to enhance the model for extracting fine-grained information from expressions and comprehending semantic information of global two-dimensional structures, multi-scale feature fusion was introduced in the encoding stage. Secondly, paired handwritten and printed mathematical expressions were introduced for training the mutual learning model, which includes decoder loss and context matching loss to learn LaTeX grammar as well as semantic invariance between handwritten and printed mathematical expressions respectively to improve the robustness of the model to different writing styles. Experimental validation was performed on the CROHME 2014/2016/2019 dataset. After introducing the multiscale feature fusion mechanism, the expression correctness rate reaches 55.25%, 52.31%, 53.72%, respectively. After introducing the mutual learning mechanism, the expression correct rate reaches 55.43%, 53.53%, 53.79%, respectively. The expression correctness rate reaches 58.88%, 55.10%, 57.05% after introducing both mechanisms at the same time. It is proved experimentally that the proposed method can effectively extract the features in formulas at different scales and overcome the problems of different handwriting styles and small amount of data by mutual learning mechanism. In addition, the experimental results on the HME100K dataset verified the effectiveness of the proposed model.  
摘要:  
脱机手写数学公式二维结构复杂,其中字符多变的尺度以及书写风格的变换不一都会增大手写数学公式识别的难度。文中提出了一个基于多尺度特征融合的互学习模型。首先,在编码阶段引入了多尺度特征融合的方式改进模型,以提升模型对公式中细粒度信息的提取能力以及加强对全局二维结构的语义信息理解;其次,引入了成对的手写体、打印体数据来进行互学习模型的训练,该模型包括解码器损失和上下文匹配损失,分别学习LaTeX语法以及手写体、打印体之间的语义不变性,提高模型对不同书写风格的鲁棒性,提升对公式整体信息的理解能力。在CROHME 2014/2016/2019数据集上进行实验验证,结果发现:引入多尺度特征融合机制后,表达式正确率分别达到55.25%、 52.31%、53.72%;引入互学习机制后,表达式正确率分别达到55.43%、53.53%、 53.79%;同时引入两种机制后,表达式正确率分别达到58.88%、55.10%、57.05%。经实验证明,文中提出的方法能够有效提取公式中不同尺度下的特征,并通过互学习机制克服手写风格不一、数据量少等问题。此外,在HME100K数据集上的实验结果也验证了文中提出模型的有效性。  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1000-565X  
入藏号:  
CSCD:7669169  
Indexed Date:  
2024-05-17  
Institution Address:  
北京工业大学信息学部, 北京 100124, 中国  
Institution Address:  
Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China  
Institution:  
北京工业大学信息学部  
Institution:  
Faculty of Information Technology, Beijing University of Technology

### 210. Remote sensing image semantic segmentation network based on multi-scale

摘要: Semantic segmentation is a crucial method for recognizing and classifying objects in high-resolution remote sensing images (HRRSIs). However, due to the problems of varying target scale and difficulty in determining the edges of small-scale targets in remote sensing images, traditional semantic segmentation models perform poorly. To address this issue, we propose a multi-scale feature enhancement network (MFENet) to improve the segmentation accuracy of small-scale objects in HRRSIs. MFENet considers the differences between objects of different scales and selects more suitable receptive fields to enhance the extraction of multi-scale semantic features. We propose a composite atrous multi-scale feature fusion (CAMFF) module to enhance the extraction of spatial detail and semantic information of features at different scales. In addition, we propose an improved composite atrous spatial pyramid pooling (C-ASPP) module to enhance the network feature extraction capability across multiple scales. We also propose a network structure that combines the C-ASPP module with the efficient channel attention (ECA) module in parallel, which performs better to extract contextual information. Our experimental evaluations on the Potsdam and Vaihingen datasets demonstrate the effectiveness of our Network, It F1 score reaching 93.33% and 94.66% respectively.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1010-6049  
eISSN:  
1752-0762  
入藏号:  
WOS:001264700000001  
Indexed Date:  
2024-07-16  
Institution Address:  
Qinghai Univ, Dept Comp Technol & Applicat, Xining, Qinghai, Peoples R China  
Institution Address:  
Qinghai Univ, Qinghai Key Lab Media Integrat Technol & Commun, Xining, Qinghai, Peoples R China  
Institution:  
Institution:

### 211. A Dense Encoder-Decoder Network with Feedback Connections for

摘要: To meet the need for multispectral images having high spatial resolution in practical applications, we propose a dense encoder-decoder network with feedback connections for pan-sharpening. Our network consists of four parts. The first part consists of two identical subnetworks, one each to extract features from PAN and MS images, respectively. The second part is an efficient feature-extraction block. We hope that the network can focus on features at different scales, so we propose innovative multiscale feature-extraction blocks that fully extract effective features from networks of various depths and widths by using three multiscale feature-extraction blocks and two long-jump connections. The third part is the feature fusion and recovery network. We are inspired by the work on U-Net network improvements to propose a brand new encoder network structure with dense connections that improves network performance through effective connections to encoders and decoders at different scales. The fourth part is a continuous feedback connection operation with overfeedback to refine shallow features, which enables the network to obtain better reconstruction capabilities earlier. To demonstrate the effectiveness of our method, we performed several experiments. Experiments on various satellite datasets show that the proposed method outperforms existing methods. Our results show significant improvements over those from other models in terms of the multiple-target index values used to measure the spectral quality and spatial details of the generated images.  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:000726538500001  
Indexed Date:  
2021-12-11  
Institution Address:  
Chongqing Univ Posts & Telecommun, Coll Comp Sci & Technol, Chongqing 400065, Peoples R China  
Institution:

### 212. Focus on Point: Parallel Multiscale Feature Aggregation for Lane Key

摘要: Lane detection, as a basic environmental perception task, plays a significant role in the safety of automatic driving. Modern lane detection methods have obtained a better performance in most scenarios, but many are unsatisfactory in various scenarios, with a weak appearance (e.g., serious vehicle occlusion, dark shadows, ambiguous markings, etc.), and have issues in simplifying model predictions and flexibly detecting lanes of a non-fixed structure and number. In this work, we abstracted the lane lines as a series of discrete key points and proposed a lane detection method of parallel multi-scale feature aggregation based on key points, FPLane. The main task of FPLane is to focus on the precise location of key points in the global lanes and aggregate the global detection results into the local geometric modeling of lane lines by using the idea of association embedding. Furthermore, this work proposes the parallel Multi-scale Feature Aggregation network (MFANet) in FPLane integrating the context information of multi-scale feature mappings to take full advantage of the prior information of adjacent positions. In addition, MFANet incorporates the Double-headed Attention Feature Fusion Up-sampling module, which can facilitate the network to accurately recognize and detect objects under extreme scale variation. Finally, our method is tested on Tusimple and CULane lane detection datasets; the results show that the proposed method outperforms the current mainstream methods: the accuracy and F1-score of the model are 96.82% and 75.6%, respectively, and the real-time detection efficiency of the model can maintain 28 ms.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:000818273200001  
Indexed Date:  
2022-07-06  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Elect & Informat Engn, Nanjing 210044, Peoples R China  
Institution:

### 213. A Multiscale and Multilayer Feature Extraction Network With Dual

摘要: A tropical cyclone (TC) is a type of catastrophic weather encountered in the tropical or subtropical ocean, and it is of great significance to accurately estimate its intensity. Many estimation methods based on statistics have been proposed, but these methods have obvious problems, such as poor robustness and low accuracy. Therefore, in this study, a TC intensity estimation method is proposed based on satellite image data using the Xception network as a backbone. The main idea of the proposed method is to estimate the TC maximum wind speed by image feature extraction. First, a Laplacian pyramid image fusion method for the infrared (IR) and water vapor (WV) channels of satellite images is adopted to enhance the total amount of information in the basic input data of the model. Second, an optimization strategy for the depth and width of the Xception network model is proposed with the objective of reducing parameter redundancy and improving the estimation accuracy. Third, a multiscale feature extraction module and a multilayer feature fusion module are designed to realize the fusion of different features. In addition, a dual attention module is introduced to allow the model to focus on the key regions of cyclone images. Finally, the proposed network is evaluated on the HURSAT, FY-2, and Gridsat datasets. The results show that, on average, the maximum cyclone wind speed estimation errors, the mean absolute error (MAE), and root mean square error (RMSE), are 8.0% and 11.4% lower than the state-of-the-art models on the three datasets.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001166583100027  
Indexed Date:  
2024-03-24  
Institution Address:  
Zhejiang Univ, Inst Marine Sensing & Networking, Zhoushan 316021, Peoples R China  
Institution Address:  
Zhejiang Univ, Coll Elect Engn, Hangzhou 310027, Peoples R China  
Institution Address:  
Zhejiang Univ, Inst Marine Sensing & Networking, Ocean Coll, Zhoushan 316021, Peoples R China  
Institution Address:  
Zhejiang Univ, Key Lab Ocean Observat Imaging Testbed, Zhoushan 316021, Peoples R China  
Institution Address:  
Pengcheng Lab, Res Ctr Artificial Intelligence, Shenzhen 518055, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Pengcheng Lab

### 214. Improved YOLO-v5 model for boosting face mask recognition accuracy on

摘要: The World Health Organization (WHO) claimed that wearing masks can effectively reduce the spread of COVID-19. In this paper, we devise novel YOLO-v5s detectors for improving the accuracy of recognizing face masks under heterogeneous IoT computing platforms. First, a module of coordinate attention is merged with the original YOLO-v5s backbone to enhance the model's attention to key information and filter out redundant information. Second, the original feature pyramid network (FPN) in the neck network is replaced with the bidirectional FPN for fast multiscale feature fusion. Third, the block of adaptive spatial feature fusion is embedded into the head layer to further solve the problem of feature inconsistency with different scales. Last, the SCYLLA-IoU metric serves as the new bounding box loss function to accelerate the model convergence. The training results show that applying the improved YOLO-v5s model to the AIZOO dataset yields an improvement in recognition accuracy of +2.2% and that the improved model has lower false detections and better detection of dense crowds. In addition, we propose a new evaluation standard of the scale limit factor to measure the detection performance of the model with different shooting distances. The proposed model is deployed on three heterogeneous IoT platforms, the Google Colab cloud, a personal computing terminal, and an Nvidia Jetson Nano edge computing device, to validate the inference feasibility, effectiveness, and robustness.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
9  
ISSN:  
2543-1536  
eISSN:  
2542-6605  
入藏号:  
WOS:001052430300001  
Indexed Date:  
2023-08-31  
Institution Address:  
Hubei Normal Univ, Sch Elect Engn & Automat, 11 Cihu Rd, Huangshi 435002, Hubei, Peoples R China  
Institution:

### 215. WSMsFNet: Joint the Whole Supervision and Multiscale Fusion Network for

摘要: Remote sensing image change detection aims to extract high-level semantic feature to identify the changed areas (CAs) between dual-temporal images (DTIs). However, the diversity in the CA shape and size poses certain challenge to the change detection (CD) task. Besides, different illumination conditions in the same scene of the DTI further increase the CD difficulty. In response to these above issues, this article proposes a multiscale feature fusion CD network-WSMsFNet, which fully utilizes the local and global information of multiscale features to achieve comprehensive representation of the change scene. In addition, the network improves the feature extraction ability of each module through the whole process supervision loss function. First, the network hierarchically extracts different scale information of the two temporal RS. Then, special information enhancement and fusion modules are constructed for various feature layers (i.e., the same level, adjacent level, and global features), aiming to enhance the local feature representation ability and contextual information relevance of the deep network. Finally, the whole-process loss function is set to supervise the intermediate layer learning, which can effectively enhance the feature representation ability and guide feature extraction direction of each module. Experiments have shown that the WSMsFNet has achieved significant results in both qualitative and quantitative indicators.  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001301018500019  
Indexed Date:  
2024-09-07  
Institution Address:  
North Univ China, Dept Comp Sci & Technol, Taiyuan 030051, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Geog Sci & Nat Resources Res, State Key Lab Resources & Environm Informat Syst, Beijing 100101, Peoples R China  
Institution:  
Institution:

### 216. Improved YOLOv3 Algorithm for Multi-target Detection of Traffic

摘要: A YOLOv3-based multi-target detection method is proposed to address the problem of missed and false detection caused by the small percentage of target pixels and mutual occlusion in traffic scenes. The method implants a spatial pyramid pooling module in the YOLOv3 network structure to enhance feature representation, and a multi-scale feature fusion mechanism is proposed to obtain both spatial information and semantic information. The semantic information of the target to be detected is refined by extending the prediction branch of the prediction layer. In addition, the improved K-means++ clustering algorithm is used to extract the initial center of the priori box and improve the matching degree between the prediction anchor box and the target to be detected. Meanwhile, a flexible non-maximum suppression algorithm is applied to adjust the confidence score flexibly. The experimental results based on the hybrid data set show that the proposed method improves the detection accuracy effectively.  
摘要:  
针对交通场景中目标像素占比小、互相遮掩等因素造成漏检、误检的问题,提出了基于YOLOv3的多目标检测方法。该方法在YOLOv3网络结构中植入空间金字塔池化模块以增强特征表达,同时提出一种多尺度特征融合机制兼顾获取空间信息和语义信息,通过扩展预测层的预测分支来细化待检目标的语义信息。此外,将改进的K均值聚类算法用于提取先验框的初始中心点,提升预测锚框与待检目标的匹配度,并运用柔性非极大值抑制算法进行置信分数的灵活调整。基于混合数据集的实验结果表明,所提方法有效地提升了检测精度。  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1007-5321  
入藏号:  
CSCD:7335825  
Indexed Date:  
2023-03-03  
Institution Address:  
兰州交通大学机电技术研究所, 兰州, 甘肃 730070, 中国  
Institution Address:  
Mechatronics T&R Institute, Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China  
Institution:  
兰州交通大学机电技术研究所  
Institution:  
Mechatronics T&R Institute, Lanzhou Jiaotong University

### 217. Traffic target detection based on context enhancement and feature

摘要: The multiscale feature fusion strategy has made substantial strides in object detection; however, it may result in feature loss or boundary distortion for larger objects when optimized for small object detection. In order to capture additional contextual information, this paper suggests a novel feature pyramid composite network structure that improves object feature extraction by incorporating residual feature extraction. In order to mitigate feature loss during the fusion process, a unified module is implemented to collect and combine global data. Furthermore, in order to optimize inter-layer information flow and mitigate interference from conflicting information, feature refinement is implemented in both channel and spatial dimensions. This method enhances the detection performance of small objects and maintains the critical features of larger objects. Experimental results indicate that this method surpasses other object detection models in terms of detection precision and achieves a 2.3% higher detection accuracy on the KITTI dataset than YOLOv9.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0920-8542  
eISSN:  
1573-0484  
入藏号:  
WOS:001405598700010  
Indexed Date:  
2025-02-01  
Institution Address:  
Xinjiang Univ, Urumqi 830047, Peoples R China  
Institution:

### 218. Object Detection of Remote Sensing Image Based on Multi-Scale Feature

摘要: In view of the small size and dense distribution of remote sensing image targets, this paper adds a detection head P2 specifically for small-scale targets on the basis of the three detection layers of the original YOLOv5 model, and involves the shallow high-resolution feature map in the subsequent multi-scale feature fusion module. The problem of losing the key feature information of the small-scale target in the process of multiple downsampling is effectively avoided. Firstly, an enhanced multi-scale feature fusion pyramid network DSI-FPN is designed. The FPN+PAN network is optimized by using DepthwiseSparable Convolution and Involution operators with fewer parameters and computations, as well as a spatial attention mechanism to generate feature graphs with richer information for network detection tasks. Secondly, we propose an adaptive channel spatial attention mechanism SCBAM, which introduces a self-attention mechanism into CBAM module to add non-local information to the interaction that originally had only local information, breaks the convolution kernel limit, expands the model receptive field, and improves the feature expression ability of the model. Thirdly, in order to solve the problem of insufficient computing power when deploying the target detector for equipment, we propose a network knowledge distillation framework for joint teachers based on the feature layer. The distillation loss of teacher is designed, and the trend of student online learning is adjusted dynamically by balancing the contributions of teacher network and truth value. The detection accuracy of the student network is obviously improved, and the parameters and model size of the network are effectively reduced. Finally, Comparing with other remote sensing image object detection methods, the experimental results show that the approach presented has better detection effect for small-scale targets of remote sensing images under different lighting conditions. The detection accuracy reached 43.9%, and 7.4% higher than that of the original model. After knowledge distillation, the model parameters are reduced to 1/3 of the original, and the detection accuracy is 40.2%.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
7  
被引频次合计:  
8  
ISSN:  
2169-3536  
入藏号:  
WOS:001145678400001  
Indexed Date:  
2024-01-26  
Institution Address:  
Harbin Univ Commerce, Sch Comp & Informat Engn, Harbin 150028, Peoples R China  
Institution Address:  
Jinan Inspur Data Technol Co Ltd, Jinan 250000, Peoples R China  
Institution:  
Institution:  
Jinan Inspur Data Technol Co Ltd

### 219. Lite-FPN for keypoint-based monocular 3D object detection

摘要: 3D object detection with a single image is an essential and challenging task for autonomous driving. Multi-scale feature fusion is effective for keypoint-based monocular 3D object detectors to boost performance within a large range of scales and distances. However, the existing FPN modules inevitably increase latency owing to the further extraction and merging operations on multi-scale feature maps. In this paper, we propose a lightweight feature pyramid network called Lite-FPN for keypoint-based monocular 3D object detectors that perform multi-scale feature fusion only at sparsely distributed keypoint locations. Besides, to alleviate the misalignment between classification score and localization precision, we propose an effective regression loss named attention loss, which assigns predictions with misaligned classification score and localization precision larger weights in the training stage. Extensive experiments based on several state-of-the-art keypoint-based detectors on the KITTI and nuScenes datasets show that our proposed methods manage to achieve significant accuracy improvements. Meanwhile, the enhanced SMOKE with our Lite-FPN module surpasses the baseline enhanced by the classic FPN over 19 FPS.(c) 2023 Elsevier B.V. All rights reserved.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
9  
ISSN:  
0950-7051  
eISSN:  
1872-7409  
入藏号:  
WOS:000989267800001  
Indexed Date:  
2023-06-05  
Institution Address:  
Tsinghua Univ, State Key Lab Automot Safety & Energy, Beijing 100084, Peoples R China  
Institution Address:  
Tsinghua Univ, Sch Vehicle & Mobil, Beijing 100084, Peoples R China  
Institution Address:  
Univ Michigan, Ann Arbor, MI USA  
Institution Address:  
Mogo Auto Intelligence & Telemetics Informat Techn, Beijing 100084, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Mogo Auto Intelligence & Telemetics Informat Techn

### 220. Serial-parallel multi-scale feature fusion for anatomy-oriented hand

摘要: Accurate hand joints detection from images is a fundamental topic that is essential for many applications in computer vision and human-computer interaction. This paper presents a two-stage network for hand joints detection from a single unmarked image by using serial-parallel multi-scale feature fusion. In stage I, the hand regions are located by an encoder-decoder network, and the features of each detected hand region are extracted by a shallow spatial hand features representation module. The extracted hand features are then fed into stage II, which consists of serially connected feature extraction modules with similar structures, called "multi-scale feature fusion" (MSFF). An MSFF contains parallel multi-scale feature extraction branches, which generate initial hand joint heatmaps. The initial heatmaps are then mutually reinforced by the anatomic relationship between hand joints. The hand joint detection accuracy shows that the proposed network overperforms the state-of-the-art methods on current datasets, 1) RHD, 2) HS, 3) MPII & NZSL, 4) DCD8-6000, with the PCK@0.2 of 0.94, 0.92, 0.84, 0.97. Meanwhile, one hand in the image takes between 24 and 37 ms to process, which is adequate for supporting many real-time applications.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000967443700001  
Indexed Date:  
2023-04-28  
Institution Address:  
Xi An Jiao Tong Univ, Sch Life Sci & Technol, Key Lab Biomed Informat Engn, Minist Educ, Xian 710049, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Bioinspired Engn & Biomech Ctr BEBC, Xian 710049, Peoples R China  
Institution Address:  
Xidian Univ, Acad Adv Interdisciplinary Res, Xian, Peoples R China  
Institution Address:  
Donghua Univ, Shanghai, Peoples R China  
Institution Address:  
Educ Univ Hong Kong, Dept Math & Informat Technol, Hong Kong, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 221. MLANet: multi-level attention network with multi-scale feature fusion

摘要: Estimating the population in a given scene is a process known as crowd counting. The field has recently garnered significant attention, and many innovative methods have emerged. However, intense scale variations and background interference make crowd counting in realistic scenes always challenging. To address these in this paper, a multi-level attention network with multi-scale feature fusion named MLANet is proposed. The network consists of three sections: a multi-level base feature extraction front-end network, a centralized dilated multi-scale feature fusion mid-end network with a global attention module, and a back-end network for the generation of density maps. By incorporating a flexible attention module and multi-scale features, the method can accurately capture crowd information at different scales and achieve accurate counting results. We evaluated the method on four public datasets (UCF\_CC\_50, ShanghaiTech, WorldExpo'10, and Beijing BRT), and the experimental results demonstrate a significant reduction in counting error when compared with existing methods.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1386-7857  
eISSN:  
1573-7543  
入藏号:  
WOS:001175257300002  
Indexed Date:  
2024-03-27  
Institution Address:  
East China Jiaotong Univ, Sch Informat Engn, Nanchang 330013, Peoples R China  
Institution Address:  
Rd Network Operat Management Co, Jiangxi Prov Commun Investment Grp Co Ltd, Nanchang 330013, Peoples R China  
Institution:  
Institution:  
Rd Network Operat Management Co

### 222. PMF-SLAM: Pose-Guided and Multiscale Feature Interaction-Based Semantic

摘要: In a dynamic environment, semantic information can assist the simultaneous localization and mapping (SLAM) system in eliminating dynamic point interference. However, most 3-D semantic segmentation methods are computationally expensive and also have low segmentation accuracy for both distant and small objects. We propose a pose-guided and multiscale feature (PMF)-SLAM method to fully exploit the interaction between 3-D semantic segmentation and SLAM and achieve efficient scene perception. The PMF-SLAM system includes three parts: a multiscale feature based segmantation network (MSF-SegNet) model, an interactive SLAM module, and a pose-guiding segmentation module. To improve the accuracy of distant and small objects, MSF-SegNet merges point-wise global features and voxel-wise local features from two branches by a designed symmetrical sparse convolution structure. In the interactive SLAM module, the coarse-to-fine registration method based on semantic information completes the estimation of the pose. To implement the interaction between segmentation and SLAM, the pose-guiding segmentation module was built to assist the segmentation thread in improving inference efficiency and ensuring segmentation consistency over time. Extensive experiments including both local experiments and the nuScenes dataset test have been conducted to validate the performance of the proposed method. Our method achieves better accuracy than multiple segmentation algorithms, significantly improving the segmentation performance of distant and small objects. The trajectory estimation accuracy is better than multiple SLAM algorithms. The code is available at https://github.com/haroldgt/MSF-SegNet.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001245613600055  
Indexed Date:  
2024-06-17  
Institution Address:  
Beijing Jiaotong Univ, State Key Lab Rail Traff Control & Safety, Beijing 100044, Peoples R China  
Institution Address:  
Beijing Jiaotong Univ, Sch Elect & Informat Engn, Beijing 100044, Peoples R China  
Institution:  
Institution:

### 223. CEDNet: A cascade encoder-decoder network for dense prediction

摘要: The prevailing methods for dense prediction tasks typically utilize a heavy classification backbone to extract multi-scale features and then fuse these features using a lightweight module. However, these methods allocate most computational resources to the classification backbone, which delays the multi-scale feature fusion and potentially leads to inadequate feature fusion. Although some methods perform feature fusion from early stages, they either fail to fully leverage high-level features to guide low-level feature learning or have complex structures, resulting in sub-optimal performance. We propose a streamlined cascade encoder-decoder network, named CEDNet, tailored for dense prediction tasks. All stages in CEDNet share the same encoder-decoder structure and perform multi-scale feature fusion within each decoder, thereby enhancing the effectiveness of multi-scale feature fusion. We explored three well-known encoder-decoder structures: Hourglass, UNet, and FPN, all of which yielded promising results. Experiments on various dense prediction tasks demonstrated the effectiveness of our method.1  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:001350104300001  
Indexed Date:  
2024-11-12  
Institution Address:  
Tsinghua Univ, Inst AI, McGovern Inst Brain Res, Tsinghua Lab Brain & Intelligence THBI,IDG,Bosch J, Beijing 100084, Peoples R China  
Institution Address:  
Huazhong Univ Sci & Technol, Sch Artificial Intelligence & Automat, Wuhan 430074, Peoples R China  
Institution Address:  
Chinese Inst Brain Res CIBR, Beijing 100010, Peoples R China  
Institution:  
Institution:  
Institution:

### 224. Video Super-Resolution Using Multi-Scale and Non-Local Feature Fusion

摘要: Video super-resolution can generate corresponding to high-resolution video frames from a plurality of low-resolution video frames which have rich details and temporally consistency. Most current methods use two-level structure to reconstruct video frames by combining optical flow network and super-resolution network, but this process does not deeply mine the effective information contained in video frames. Therefore, we propose a video super-resolution method that combines non-local features and multi-scale features to extract more in-depth effective information contained in video frames. Our method obtains long-distance effective information by calculating the similarity between any two pixels in the video frame through the non-local module, extracts the local information covered by different scale convolution cores through the multi-scale feature fusion module, and fully fuses feature information using different connection modes of convolution cores. Experiments on different data sets show that the proposed method is superior to the existing methods in quality and quantity.  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
2079-9292  
入藏号:  
WOS:000794699400001  
Indexed Date:  
2022-05-23  
Institution Address:  
Xian Univ Technol, Fac Automat & Informat Engn, Xian 710048, Peoples R China  
Institution Address:  
Xian Univ Technol, Sch Printing Packaging & Digital Media, Xian 710054, Peoples R China  
Institution Address:  
Hubei Univ Automot Technol, Sch Elect & Informat Engn, Shiyan 442002, Peoples R China  
Institution:  
Institution:  
Institution:

### 225. A multi-scale feature fusion spatial-channel attention model for

摘要: Background subtraction is an essential task in computer vision, and is often used as a pre-processing step for many advanced tasks. In this work, we propose a novel multi-scale feature fusion attention mechanism network to tackle cross-scene background subtraction. The cross-fusion of feature maps at different stages of the encoder makes the features input into the decoder contain low-level and high-level information. The spatial-channel attention based on the weight matrix makes the model focus on processing information related to foreground extraction. We evaluate the proposed model on the CDnet-2014 dataset with two scene-independent evaluation strategies and obtain competitive F-Measure. In addition, to evaluate the generalization ability of the model, we perform a cross-dataset evaluation scheme on the LASIESTA and SBI2015 datasets. The overall F-Measure of the model is 0.89 and 0.93, respectively. Experimental results demonstrate that the model performs well compared to the current state-of-the-art methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0942-4962  
eISSN:  
1432-1882  
入藏号:  
WOS:001040551300002  
Indexed Date:  
2023-08-12  
Institution Address:  
Hefei Univ Technol, Sch Microelect, Hefei 230009, Peoples R China  
Institution:

### 226. Detection of X-Ray Contraband by Adaptive and Multi-Scale Feature Fusion

摘要: To resolve the problems of spatial multi-scale variation, background interference and model complex of X-ray security inspection contraband images, a lightweight YOLOv5 model with spatial adaptation and multi-scale feature fusion is proposed. Taking YOLOv5 as the basic framework, the adaptive spatial feature fusion mechanism is introduced to suppress the influence of feature scale differences, and the bidirectional feature weighted fusion is integrated with the bidirectional feature pyramid network; the lightweight channel attention mechanism is used to obtain accurate position information and enhance the expression of effective features. Meanwhile, GhostConv is used to replace part of Conv to reduce the computational complexity of the network. This model achieves mAP of 94.2%, 92.8% and 83.3% on three public datasets such as OPIXray, SIXray and HiXray, respectively, which is 5.4, 0.5 and 1.7 percentage points higher than the baseline model. And the model training time is not significantly increased. It takes into account the accuracy and speed of model detection, which is superior to many current advanced algorithms.  
摘要:  
针对X光安检违禁品图像空间多尺度变化、背景干扰及模型复杂等问题,提出了空间自适应与多尺度特征融合的YOLOv5轻量模型。以YOLOv5为基本框架,引入自适应空间特征融合机制抑制特征尺度差异的影响,结合双向特征金字塔网络集成了特征双向加权融合;采用轻量化通道注意力机制获得编码的位置信息,增强有效特征的表达;同时利用GhostConv替换部分Conv降低网络计算复杂度。此模型在OPIXray、SIXray、HiXray等3个公开数据集上mAP分别达到94.2%、92.8%、83.3%,比基线模型分别提高了5.4、0.5、1.7个百分点,且未明显改变推理效率,较好兼顾了模型检测精度与速度,优于当前诸多先进算法。  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7657618  
Indexed Date:  
2024-04-26  
Institution Address:  
中国人民公安大学信息网络安全学院, 北京 102600, 中国  
Institution Address:  
中国人民公安大学信息网络安全学院;;公安部安全防范技术与风险评估重点实验室, ;;公安部安全防范技术与风险评估重点实验室, ;;, 北京;;北京 102600;;102600, 中国  
Institution Address:  
College of Information and Cyber Security,People's Public Security University of China, Beijing 102600, China  
Institution Address:  
College of Information and Cyber Security,People's Public Security University of China;;Key Laboratory of Security Technology and Risk Assessment Ministry of Public Security, ;;Key Laboratory of Security Technology and Risk Assessment Ministry of Public Security, ;;, Beijing;;Beijing 102600;;102600  
Institution:  
中国人民公安大学信息网络安全学院  
Institution:  
中国人民公安大学信息网络安全学院;;公安部安全防范技术与风险评估重点实验室  
Institution:  
College of Information and Cyber Security,People's Public Security University of China  
Institution:  
College of Information and Cyber Security,People's Public Security University of China;;Key Laboratory of Security Technology and Risk Assessment Ministry of Public Security

### 227. Monocular Depth Estimation With Multi-Scale Feature Fusion

摘要: Depth estimation from a single image is a crucial but challenging task for reconstructing 3D structures and inferring scene geometry. However, most existing methods fail to extract more detailed information and estimate the distant small-scale objects well. In this paper, we propose a monocular depth estimation based on multi-scale feature fusion. Specifically, to obtain input features of different scales, we first feed the input images of different scales to pre-trained residual networks with sharing weights. Then, an attention mechanism is used to learn the salient features at different scales, which can integrate detailed information at large scale feature maps and scene information at small scale feature maps. Furthermore, inspired by the dense atrous spatial pyramid pooling in semantic segmentation, we build a multi-scale feature fusion dense pyramid to further improve the ability of the feature extraction. Last, a scale-invariant error loss is used to predict depth maps in log space. We evaluate our method on several public benchmark datasets (including NYU Depth V2 and KITTI). The experiment results show that the proposed method obtains better performance than the existing methods and achieves state-of-the-art results.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
16  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
17  
ISSN:  
1070-9908  
eISSN:  
1558-2361  
入藏号:  
WOS:000641957600010  
Indexed Date:  
2021-05-20  
Institution Address:  
Dalian Univ Technol, Sch Informat & Commun Engn, Dalian 116023, Peoples R China  
Institution:

### 228. LLR-MVSNet: a lightweight network for low-texture scene reconstruction

摘要: In recent years, learning-based MVS methods have achieved excellent performance compared with traditional methods. However, these methods still have notable shortcomings, such as the low efficiency of traditional convolutional networks and simple feature fusion, which lead to incomplete reconstruction. In this research, we propose a lightweight network for low-texture scene reconstruction (LLR-MVSNet). To improve accuracy and efficiency, a lightweight network is proposed, including a multi-scale feature extraction module and a weighted feature fusion module. The multi-scale feature extraction module uses depth-separable convolution and point-wise convolution to replace traditional convolution, which can reduce network parameters and improve the model efficiency. In order to improve the fusion accuracy, a weighted feature fusion module is proposed, which can selectively emphasize features, suppress useless information and improve the fusion accuracy. With rapid computational speed and high performance, our method surpasses the state-of-the-art benchmarks and performs well on the DTU and the Tanks & Temples datasets. The code of our method will be made available at https://github.com/wln19/LLR-MVSNet.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0942-4962  
eISSN:  
1432-1882  
入藏号:  
WOS:001309450700002  
Indexed Date:  
2024-09-15  
Institution Address:  
Nanjing Univ, Jiangsu Prov Key Lab Geog Informat Sci & Technol, Key Lab Land Satellite Remote Sensing Applicat, Sch Geog & Ocean Sci,Minist Nat Resources, Nanjing 210023, Peoples R China  
Institution Address:  
Jiangsu Ctr Collaborat Innovat Novel Software Tech, Nanjing 210023, Jiangsu, Peoples R China  
Institution:  
Institution:  
Jiangsu Ctr Collaborat Innovat Novel Software Tech

### 229. 3D UNeXt:lightweight and efficient network for effective brain

摘要: In order to solve the drawbacks of existing brain extraction network,i.e.,complex network structure,large amounts of parameters and low inference speed,this paper proposed a novel network 3D UNeXt for fast and effective brain extraction.3D UNeXt greatly reduced parameters and the number of floating point operators(FLOPs),and achieved promising results with the combination of 3D convolution,3D MLP and multi-scale feature fusion.3D UNeXt used U-Net as the basic architecture and employed 3D convolutional modules to obtain local features in encoding stage.Specifically,the proposed 3D MLP module at the bottleneck stage enhanced the extraction of global features and long-range dependencies among them.In decoding stage,this paper designed a lightweight multiscale feature fusion module to effectively fuse multiscale low-level features and high-level counterparts.In detail,the 3D MLP module performed linear shift operations in three different axes to obtain global receptive fields from different dimension features and establish long-range dependencies among them.This paper compared 3D UNeXt with other counterparts on three datasets:IBSR,NFBS,and HTU-BrainMask.Experimental results show that the 3D UNeXt is superior over other baselines in terms of network parameters,FLOPs,inference accuracy,and inference speed.  
摘要:  
为了解决现有脑提取网络结构复杂、参数量大且推理速度不高的问题,受UNeXt启发,提出一种基于3D卷积、3D多层感知机(multilayer perception,MLP)和多尺度特征融合的轻量级快速脑提取网络3D UNeXt,极大地减少了参数和浮点运算量,取得了令人满意的结果。3D UNeXt以U-Net为基本架构,在编码阶段使用3D卷积模块获取局部特征;在瓶颈阶段通过3D MLP模块获取全局特征和特征之间的远程依赖;在解码阶段借助多尺度特征融合模块高效融合浅层特征和深层特征。特别地,3D MLP模块在三个不同特征轴向进行线性移位操作,以获取不同维度特征的全局感受野并建立它们之间的远程依赖。在IBSR、NFBS和HTU-BrainMask三个数据集上进行实验,以和先进网络进行对比。实验结果表明,3D UNeXt在网络参数、浮点运算量、推理精度和速度等方面显著优于现有模型。  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1001-3695  
入藏号:  
CSCD:7752826  
Indexed Date:  
2024-08-30  
Institution Address:  
河南师范大学计算机与信息工程学院;;河南省教育人工智能与个性化学习重点实验室;;教学资源与教育质量评估大数据河南省工程实验室, ;;河南省教育人工智能与个性化学习重点实验室;;教学资源与教育质量评估大数据河南省工程实验室, 新乡;;新乡;;新乡, 河南;;河南;;河南 453007;;453007;;453007, 中国  
Institution Address:  
河南师范大学计算机与信息工程学院, 新乡, 河南 453007, 中国  
Institution Address:  
School of Computer & Information Engineering,Henan Normal University;;Key Laboratory of Artificial Intelligence & Personalized Learning in Education of Henan Province;;Big Data Engineering Laboratory of Teaching Resources & Assessment of Education Quality of Henan Province, ;;Key Laboratory of Artificial Intelligence & Personalized Learning in Education of Henan Province;;Big Data Engineering Laboratory of Teaching Resources & Assessment of Education Quality of Henan Province, Xinxiang;;Xinxiang;;Xinxiang, Henan;;Henan;;Henan 453007;;453007;;453007  
Institution Address:  
School of Computer & Information Engineering,Henan Normal University, Xinxiang, Henan 453007, China  
Institution:  
河南师范大学计算机与信息工程学院;;河南省教育人工智能与个性化学习重点实验室;;教学资源与教育质量评估大数据河南省工程实验室  
Institution:  
河南师范大学计算机与信息工程学院  
Institution:  
School of Computer & Information Engineering,Henan Normal University;;Key Laboratory of Artificial Intelligence & Personalized Learning in Education of Henan Province;;Big Data Engineering Laboratory of Teaching Resources & Assessment of Education Quality of Henan Province  
Institution:  
School of Computer & Information Engineering,Henan Normal University

### 230. EFFD: An Unsupervised Surface Defect Detection Method Based on

摘要: Deep feature reconstruction (DFR) as an unsupervised surface defect detection framework, first combined the idea of feature embedding with reconstruction. However, it concatenated all features for reconstruction, which may cause the detection results to be disturbed by some features that are not sensitive to anomalies, causing missed detection. Therefore, we made a series of improvements to DFR and proposed a method called estimation and fusion of normal sample feature distribution (EFFD), as well as its upgraded version EFFD+. Our improvements are based on the assumption that features closer to the normal sample feature distribution are more sensitive to anomalies. We used multivariate Gaussian distribution estimation (MGDE) and Bayesian multivariate Gaussian distribution fusion (BMGDF) modules to estimate and fuse the normal sample feature distribution across all levels at each scale, and made the features fused by iterative attention feature fusion (IAFF) module follow to the fused distribution of the corresponding scale. This allowed the parts of the features that are sensitive to anomalies to be assigned higher weights. In EFFD+, we proposed IAFF+, which added more branches of different receptive fields based on IAFF and used weighted multiscale feature fusion (WMSFF) module to assign the optimal weights to the features at each scale. EFFD+ is more conducive to the detection of various categories of defects. On MVTecAD, BTAD, and VisA datasets, our proposed method achieved 98.5%, 97.9%, and 98.8% average pixel-level area under the receiver operating characteristic curve (P-AUC-ROC), respectively, and 94.6%, 81.9%, and 93.1% average area under the per region overlap curve (AUC-PRO), respectively, which has reached the current advanced level. Code is available at: https://github.com/Jalexdalv/effd.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001389581900010  
Indexed Date:  
2025-01-11  
Institution Address:  
Bohai Univ, Dept Control Sci & Engn, Jinzhou 121013, Liaoning, Peoples R China  
Institution:

### 231. CA-YOLOv5 for Crowded Pedestrian Detection

摘要: Aiming at the problem of high miss-detection rate and insufficient feature fusion of YOLOv5 in crowded pedestrian detection task, the CA-YOLOv5 pedestrian detection algorithm is proposed. To solve the problem of insufficient fine-grained feature fusion in the original backbone network, Res2Block is used to rebuild the backbone network of YOLOv5, so as to improve the fine-grained feature fusion ability of the network and improve the detection accuracy. For the large change of target scale in dataset, coordinate attention is introduced to enhance the receptive field and the model's ability to accurately locate the target. Aiming at the problem that FPN structure reduces the multi-scale feature expression ability during feature fusion, the feature enhancement module is proposed to enhance the multi-scale feature expression ability. Through the structural re-parameterization method to reduce the number of parameters and computation in the model, and speed up the detection. Aiming at the common problem of crowded pedestrians in pedestrian detection task, EViT is proposed to enhance the ability of the model to pay attention to local information and improve the detection accuracy. Experimental results show that in the crowded pedestrian detection task, the detection accuracy of CA-YOLOv5 reaches 84.86%, 3.75% higher than the original algorithm, and the detection speed can reach 51 FPS, which has good detection accuracy and real-time. Therefore, it can be better applied to real-time pedestrian detection task.  
摘要:  
针对YOLOv5在拥挤行人检测任务中漏检率高、特征融合不充分等问题,提出了CA-YOLOv5行人检测算法。针对原主干网络对细粒度特征融合不充分的问题,采用Res2Block重建YOLOv5的主干网络,以提升网络的细粒度特征融合能力,提高检测精度。针对数据集目标尺度变化大的问题,引入coordinate attention(CA)模块增强感受野,增强模型对目标的精确定位能力。针对FPN结构在特征融合时导致多尺度特征表达能力下降的问题,提出特征增强模块,以增强多尺度特征的表达能力。通过结构重参数化的方法减少模型的计算量与参数量,加快目标检测速度。针对行人检测任务中普遍存在的拥挤行人问题,提出EViT模块,增强模型关注局部信息的能力,提高检测精度。实验证明,在拥挤行人检测任务中,CA-YOLOv5的检测精度达到84.86%,相较于原算法提高了3.75%,检测速度可以达到51 FPS,具有较好的检测精度与实时性。因此,CA-YOLOv5可以更好地应用于实时行人检测任务中。  
在中国科学引文数据库中的被引频次:  
5  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
12  
ISSN:  
1002-8331  
入藏号:  
CSCD:7219693  
Indexed Date:  
2022-08-05  
Institution Address:  
新疆大学软件学院, 乌鲁木齐, 新疆 830046, 中国  
Institution Address:  
School of Software, Xinjiang University, Urumqi, Xinjiang 830046, China  
Institution:  
新疆大学软件学院  
Institution:  
School of Software, Xinjiang University

### 232. Wheel Tread Anomaly Detection Based on Attentional Reverse Knowledge

摘要: Wheels are an essential part of railway trains; thus, defects on the wheel tread present serious risk regarding the safety of railway trains. Due to the limited samples of wheel tread defects in practice, the corresponding supervised detection model is insufficient. To solve this problem, an unsupervised knowledge distillation anomaly detection model is proposed to detect wheel tread anomalies. Accordingly, UNet is employed to segment the tread region and reduce the influence of non- tread regions on the anomaly detection model. An attention mechanism is then added after the multiscale feature fusion to improve the ability of the student network to reconstruct normal features in the reverse knowledge distillation structure, as well as enhance the reconstruction of normal features. From the experimental results, the improved model achieves the performance indexes of 93.8% area under receiver operating characteristic curve, 82.3% precision, 95.4% recall, and 87.0% accuracy considering the railway wheel tread dataset. Compared with the original model, the detection performance of the model is improved.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001180502700021  
Indexed Date:  
2024-03-20  
Institution Address:  
Southwest Jiaotong Univ, Sch Phys Sci & Technol, Chengdu 610000, Sichuan, Peoples R China  
Institution:

### 233. Enhanced damage segmentation in RC components using pyramid Haar wavelet

摘要: Damage identification in post-earthquake reinforced concrete (RC) structures based on semantic segmentation has been recognized as a promising approach for rapid and non-contact damage localization and quantification. In damage segmentation tasks, damage regions are often set against complex backgrounds, featuring irregular geometric boundaries and intricate textures, posing significant challenges to model segmentation performance. Additionally, the absence of public datasets exacerbates these challenges, hindering advancements in this field. In this paper, a pyramid Haar wavelet downsampling attention UNet (PHA-UNet) semantic segmentation network is proposed, and a database containing 1400 images of damaged RC components (PEDRC-Dataset) with pixel-level annotations is established. In the proposed PHA-UNet, attention mechanisms, multiscale feature fusion, Haar wavelet downsampling, and transfer learning are introduced to address above challenges. Finally, the proposed PHA-UNet is compared with four existing image segmentation architectures on both the Cityspace and the PEDRC-Dataset.  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0926-5805  
eISSN:  
1872-7891  
入藏号:  
WOS:001315110200001  
Indexed Date:  
2024-09-23  
Institution Address:  
Xian Univ Architecture & Technol, Coll Civil Engn, Xian, Peoples R China  
Institution Address:  
Xian Univ Architecture & Technol, State Key Lab Green Bldg Western China, Xian, Peoples R China  
Institution Address:  
China Earthquake Adm, Inst Engn Mech, Harbin 150080, Peoples R China  
Institution:  
Institution:  
Institution:

### 234. A Single-Stage Arbitrary-Oriented Detector Based on Multiscale Feature

摘要: Ship detection is a challenging task in the synthetic aperture radar (SAR) automatic target recognition due to the large aspect ratio, arbitrary orientation, and dense arrangement of ships and severe background interference in the inshore scenes. Although considerable progress has been made in recent research, there have still been certain challenges in achieving fast and efficient detection of arbitrary-oriented ships in SAR images. To address these challenges, this article proposes a single-stage detection method based on multiscale feature fusion and calibration. The proposed detection method can detect arbitrary-oriented ships in SAR images with high accuracy and speed. Specifically, a head network with the stepwise regression from the coarse- to fine-grained detection is designed to detect arbitrary-oriented ships accurately. In addition, a feature enhancement module is constructed to fuse and refine shallow texture features and deep semantic features, aiming to obtain multiscale fusion features containing sufficient contextual information. Finally, an attention module is used to calibrate multiscale fusion features to highlight the ship information while suppressing interference from the surrounding background. The effectiveness of the proposed method is verified by experiments on two public SAR ship datasets and a panoramic SAR image. The experimental results show that compared with the other rotation detectors, the proposed detection method has competitive detection results and can achieve state-of-the-art detection performance.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
8  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000862384000002  
Indexed Date:  
2022-10-15  
Institution Address:  
Chinese Acad Sci, Dept Space Microwave Remote Sensing Syst, Aerosp Informat Res Inst, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 101408, Peoples R China  
Institution:  
Institution:

### 235. LVD-YOLO: An efficient lightweight vehicle detection model for

摘要: Vehicle detection is a fundamental component of intelligent transportation systems. However, current algorithms often encounter issues such as high computational complexity, long execution times, and significant resource demands, making them unsuitable for resource-limited environments. To overcome these challenges, we propose LVD-YOLO, a Lightweight Vehicle Detection Model based on YOLO. This model incorporates the EfficientNetv2 network structure as its backbone, which reduces parameters and enhances feature extraction capabilities. By utilizing a bidirectional feature pyramid structure along with a dual attention mechanism, we enable efficient information exchange across feature layers, thereby improving multiscale feature fusion. Additionally, we refine the model's loss function with SIoU loss to boost regression and prediction performance. Experimental results on the PASCAL VOC and MS COCO datasets show that LVD-YOLO outperforms YOLOv5s, achieving a 0.5% increase in accuracy while reducing FLOPs by 64.6% and parameters by 48.6%. These improvements highlight its effectiveness for use in resource-constrained environments.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0262-8856  
eISSN:  
1872-8138  
入藏号:  
WOS:001318150900001  
Indexed Date:  
2024-09-29  
Institution Address:  
Shandong Technol & Business Univ, Sch Informat & Elect Engn, Yantai 264005, Peoples R China  
Institution:

### 236. Method of feature pyramid and attention enhancement network for pavement

摘要: The detection of pavement crack information plays an important role in the service evaluation of infrastructure engineering performance. The randomness of road noise, pollution, and spots make it difficult to ensure the accuracy of traditional image processing algorithms to detect cracks. Therefore, the crack identification technology under complex background conditions is particularly important. We propose a network architecture, named feature pyramid and attention enhancement network (FPAEN) for pavement crack detection. Through a new multiscale feature fusion module, the deep semantic information is integrated into the low-level convolution stage layer by layer, so the context information circulates throughout the whole network. In addition, the attention enhancement mechanism and the multiscale dilated convolution module are used to strengthen the network model's ability to locate the crack pixels. To evaluate the performance of the proposed approach, we train FPAEN on the DeepCrack dataset and test it on four datasets. Experimental results show that our method outperforms other advanced crack detection approaches in F-score and mean intersection over union. (C) 2022 SPIE and IS&T  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:000818115800021  
Indexed Date:  
2022-07-06  
Institution Address:  
Chongqing Univ Posts & Telecommun, Coll Comp Sci & Technol, Chongqing, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Coll Software Engn, Chongqing, Peoples R China  
Institution:  
Institution:

### 237. DSHFNet: Dynamic Scale Hierarchical Fusion Network Based on

摘要: With the continuous improvement of satellite sensor performance, it is becoming easier to obtain different types of remote sensing (RS) data from multiple sensors, and the fusion of hyperspectral (HS) images and light detection and ranging (LiDAR) for land use/land cover (LULC) classification has become a research hotspot. However, the current mainstream methods still have defects in feature extraction and feature fusion. In the feature extraction stage, previous methods usually use a single-scale patch as input and a fixed convolution kernel for feature extraction, which makes it difficult to extract features in line with different land cover types at the same time and to obtain high-quality features. Although multiscale feature extraction can solve the one-sidedness problem of single-scale features, it also brings the challenge of high-dimensional multiscale features. In the feature fusion stage, the current fusion methods are relatively simple. Therefore, we propose a dynamic scale hierarchical fusion network (DSHFNet) for fusion classification of HS images and LiDAR data. By calculating the similarity in the scale space and judging the information at different scales through the threshold value, the appropriate scale features are dynamically selected, the small-scale features are integrated into the large-scale features, and the dimensionality of the features is reduced. This method solves the unreliability problem of single-scale features and the high-dimensional problem of multiscale features. In the feature fusion process, different attention modules are used for hierarchical fusion, spatial attention modules are used for shallow fusion and joint feature extraction, and modal attention modules are used for deep fusion of joint features and features from different sensors to achieve complete complementarity of features. Experimental evaluations on three real RS datasets demonstrate the superiority of the proposed method compared with existing methods. The source code can be downloaded at https://github.com/SYFYN0317/DSHFNet.  
Web of Science 核心合集中的 "被引频次":  
20  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
20  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001068941300001  
Indexed Date:  
2023-10-18  
Institution Address:  
Liaoning Normal Univ, Sch Geog, Dalian 116029, Peoples R China  
Institution Address:  
Liaoning Normal Univ, Sch Comp Sci & Artificial Intelligence, Dalian 116029, Peoples R China  
Institution Address:  
Liaoning Normal Univ, Sch Geog, Dalian 116029, Peoples R China  
Institution:  
Institution:  
Institution:

### 238. A Lightweight Remote Sensing Small Target Image Detection Algorithm

摘要: In response to the challenges posed by small objects in remote sensing images, such as low resolution, complex backgrounds, and severe occlusions, this paper proposes a lightweight improved model based on YOLOv8n. During the detection of small objects, the feature fusion part of the YOLOv8n algorithm retrieves relatively fewer features of small objects from the backbone network compared to large objects, resulting in low detection accuracy for small objects. To address this issue, firstly, this paper adds a dedicated small object detection layer in the feature fusion network to better integrate the features of small objects into the feature fusion part of the model. Secondly, the SSFF module is introduced to facilitate multi-scale feature fusion, enabling the model to capture more gradient paths and further improve accuracy while reducing model parameters. Finally, the HPANet structure is proposed, replacing the Path Aggregation Network with HPANet. Compared to the original YOLOv8n algorithm, the recognition accuracy of mAP@0.5 on the VisDrone data set and the AI-TOD data set has increased by 14.3% and 17.9%, respectively, while the recognition accuracy of mAP@0.5:0.95 has increased by 17.1% and 19.8%, respectively. The proposed method reduces the parameter count by 33% and the model size by 31.7% compared to the original model. Experimental results demonstrate that the proposed method can quickly and accurately identify small objects in complex backgrounds.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
被引频次合计:  
8  
eISSN:  
1424-8220  
入藏号:  
WOS:001220016000001  
PubMed ID:  
38733059  
Indexed Date:  
2024-05-22  
Institution Address:  
Changchun Univ Technol, Sch Comp Sci & Engn, Changchun 130012, Peoples R China  
Institution:

### 239. Person re-identification method based on fine-grained feature fusion and

摘要: Aiming at the problem of low accuracy of person re-identification (Re-ID) algorithm caused by occlusion, low distinctiveness of person features and unclear detail features in complex environment, we propose a Re-ID method based on fine-grained feature fusion and self-attention mechanism. First, we design a dilated non-local module (DNLM), which combines dilated convolution with the non-local module and embeds it between layers of the backbone network, enhancing the self-attention and receptive field of the model and improving the performance on occlusion tasks. Second, the fine-grained feature fusion screening module (3FSM) is improved based on the outlook attention module, which can realize adaptive feature selection and enhance the recognition ability to similar samples of the model. Finally, combined with the feature pyramid in the field of object detection, we propose a multi-scale feature fusion pyramid (MFFP) to improve the Re-ID tasks, in which we use different levels of features to perform feature enhancement. Ablation and comprehensive experiment results based on multiple datasets validate the effectiveness of our proposal. The mean Average Precision (mAP) of Market1501 and DukeMTMC-reID is 92.5 and 87.7%, and Rank-1 is 95.1 and 91.1% respectively. Compared with the current mainstream Re-ID algorithm, our method has excellent Re-ID performance.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0010-485X  
eISSN:  
1436-5057  
入藏号:  
WOS:001190454900001  
Indexed Date:  
2024-03-29  
Institution Address:  
Univ Elect Sci & Technol China, Chengdu 610057, Peoples R China  
Institution Address:  
UESTC, Shenzhen Inst Adv Study, Shenzhen 518110, Peoples R China  
Institution Address:  
Kash Inst Elect & Informat Ind, Kashi 610057, Xinjiang, Peoples R China  
Institution:  
Institution:  
Institution:  
Kash Inst Elect & Informat Ind

### 240. Improving vehicle detection accuracy in complex traffic scenes through

摘要: Vehicle detection is a fundamental task for automated driving systems. However, achieving robust performance in complex traffic scenarios remains a formidable challenge. In this paper, we propose a novel vehicle detection model that leverages contextual attention mechanisms and multi-scale feature fusion to effectively tackle the inherent challenges presented by complex scenarios, such as occlusion, truncation, and small-scale vehicle instances. The proposed model introduces a contextual attention module tailored to address vehicle occlusion, augmenting the model's reasoning ability and overall performance through the integration of global contextual information. Additionally, we introduce a Multi-Scale Feature Fusion Module to mitigate the impact of drastic changes in vehicle scales observed in dynamic traffic scenarios. Through the deployment of a dedicated multi-scale feature fusion module, our model adeptly adapts to significant size variations of vehicles in traffic scene images, thereby enhancing its capability to handle vehicles of varying sizes. Our contributions are validated through comprehensive qualitative and quantitative experiments conducted on both the KITTI dataset and the Cityscapes dataset. The experimental results demonstrate the exceptional robustness and accuracy of our proposed model. These findings provide conclusive evidence of the superior performance and effectiveness of our model in real-world applications.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:001409066100004  
Indexed Date:  
2025-02-05  
Institution Address:  
Southwest Jiaotong Univ, Sch Informat Sci & Technol, Xian Rd, Chengdu 611756, Sichuan, Peoples R China  
Institution:

### 241. SAE-CenterNet: Self-attention enhanced CenterNet for small dense object

摘要: The existing object detection models have not been sufficiently optimized for small dense objects. One of the most common solutions is to extract multi-scale features via feature pyramid network (FPN). However, the information loss of downsampling in multi-scale feature extraction will seriously affect the detection accuracy. Therefore, a dynamic attention convolution (DAC) for downsampling is developed, which can embed regional information for each single pixel. Besides, an attention fusion module (AFM) is also designed to alleviate the data inconsistency from different layers during multi-scale feature fusion. Based on these, the proposed model, SAE-CenterNet, has achieved optimal performance in the mainstream object detection models on the small dense rebar dataset. For example, with 6 FPS decreasing, the mAP50$ mAP\_{50}$, mAP75$ mAP\_{75}$, Recall50$ Recall\_{50}$ and Recall75$ Recall\_{75}$ of SAE-CenterNet is 87.3%, 57.6%, 88.6% and 67.2%, respectively, which are 8.0%, 13.5%, 8.3% and 10.4% higher than the baseline CenterNet, respectively.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
5  
ISSN:  
0013-5194  
eISSN:  
1350-911X  
入藏号:  
WOS:000920022900001  
Indexed Date:  
2023-02-16  
Institution Address:  
Wuhan Inst Technol, Hubei Key Lab Intelligent Robot, Wuhan, Peoples R China  
Institution:

### 242. MS2Net: Multi-Scale and Multi-Stage Feature Fusion for Blurred Image

摘要: At present, most mainstream algorithms for single image super-resolution (SISR) assume the image degradation process as an ideal degradation process (e.g. bicubic downscaling), which violates the actual degeneration conditions. In real-world image capturing, objects often move in a dynamic environment, and camera shake also often occurs, which results in serious blurs. Our work focuses on the task of image super-resolution with heavy motion blur, for which we adopt a network with two branches: one branch for image deblurring and the other one for super-resolution. Since the features obtained by the deblurring are rich in details, we apply their features as supplementary information to the super-resolution branch. Based on the adopted dual-branch framework, our major technical novelties lie in two novel modules: Multi-Scale Feature Fusion (MSFF1) module which fuses features of different scale from the deblurring branch to get local and global information, and Multi-Stage Feature Fusion (MSFF2) module which further filters useful information with attention. We evaluate the proposed method under various blur scenarios on the benchmark datasets, demonstrating competitive performance against existing methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
4  
Web of Science 核心合集中的 "被引频次":  
28  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
29  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:000835828500020  
Indexed Date:  
2022-08-13  
Institution Address:  
Northwestern Polytech Univ, Sch Comp Sci, Xian 710072, Peoples R China  
Institution Address:  
Natl Engn Lab Integrated AeroSp Ground Ocean Big, Xian 710072, Peoples R China  
Institution Address:  
Korea Adv Inst Sci & Technol, Sch Elect Engn, Daejeon 34141, South Korea  
Institution Address:  
Korea Adv Inst Sci & Technol, Robot & Comp Vis Lab, Daejeon 34141, South Korea  
Institution Address:  
Northwestern Polytech Univ, Sch Astronaut, Xian 710072, Peoples R China  
Institution:  
Institution:  
Natl Engn Lab Integrated AeroSp Ground Ocean Big  
Institution:  
Institution:  
Institution:

### 243. Adaptive multi-scale feature fusion with spatial translation for

摘要: In image segmentation tasks, contextual information is crucial as it provides essential semantic details. Multi-scale feature extraction methods help models capture this contextual information comprehensively, but they can introduce redundancy and insufficient receptive fields in some areas, particularly with large objects or complex scenes. To address these issues, we propose the Adaptive Feature Perception Module (AFPM). Inspired by the visual system, we combine the pyramid model with dilated convolutions and incorporate a spatial shift mechanism for extensive information capture.This module adaptively adjusts its focus and perception range to maximize target feature capture.Meanwhile, we introduce the Channel and Spectral Attention Module(CSAM) to model dependencies between channels and spectral domains,enabling the network to learn more discriminative features and improve segmentation accuracy. Based on these enhancements,we propose a new network model called AMFFNet. We validated its effectiveness by comparing it with several state-of-the-art methods on the PASCAL VOC 2012, Cityscapes and ADE20K datasets. The results demonstrate that AMFFNet offers superior performance.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001286387000001  
Indexed Date:  
2024-08-13  
Institution Address:  
Harbin Engn Univ, Coll Informat & Commun Engn, Harbin 150001, Heilongjiang, Peoples R China  
Institution Address:  
Harbin Engn Univ, Key Lab Adv Ship Commun & Informat Technol, Harbin 150001, Heilongjiang, Peoples R China  
Institution:  
Institution:

### 244. Ship Detection and Recognition Based on Improved YOLOv7

摘要: In this paper, an advanced YOLOv7 model is proposed to tackle the challenges associated with ship detection and recognition tasks, such as the irregular shapes and varying sizes of ships. The improved model replaces the fixed anchor boxes utilized in conventional YOLOv7 models with a set of more suitable anchor boxes specifically designed based on the size distribution of ships in the dataset. This paper also introduces a novel multiscale feature fusion module, which comprises Path Aggregation Network (PAN) modules, enabling the efficient capture of ship features across different scales. Furthermore, data preprocessing is enhanced through the application of data augmentation techniques, including random rotation, scaling, and cropping, which serve to bolster data diversity and robustness. The distribution of positive and negative samples in the dataset is balanced using random sampling, ensuring a more accurate representation of real-world scenarios. Comprehensive experimental results demonstrate that the proposed method significantly outperforms existing state-of-the-art approaches in terms of both detection accuracy and robustness, highlighting the potential of the improved YOLOv7 model for practical applications in the maritime domain.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
19  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
19  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001020942100029  
Indexed Date:  
2023-07-24  
Institution Address:  
Hainan Univ, Sch Informat & Commun Engn, Haikou 570228, Peoples R China  
Institution Address:  
Hainan Univ, Sch Cyberspace Secur, Haikou 570228, Peoples R China  
Institution Address:  
Hainan Univ, Sch Comp Sci & Technol, Haikou 570228, Peoples R China  
Institution:  
Institution:  
Institution:

### 245. Multi-scale fusion and adaptively attentive generative adversarial

摘要: The quality of images taken on rainy days is decreased due to rain streaks. These degraded images affect the performance of vision applications (e.g., face detection and verification). Besides, because of multi-directions, multi-sizes and multi-densities of rain streaks, the existing rain removal methods lack deeper insight into the rain image and are not ideal to remove rain. To address the problem, we propose a novel rain removal model named Multi-scale Fusion and Adaptively Attentive Generative Adversarial Network (MFAA-GAN) to efficiently remove rain streaks. First, in MFAA-GAN, to extract multiscale features and the correlation information between cross-scale features, we design a multiscale feature fusion module that takes two parallel residual dense blocks with different sizes of convolution kernels. Secondly, we add an adaptive attention algorithm including spatial attention and channel attention in generator to capture local and global position information respectively. Third, in the training procedure, we propose a new multi-scale perceptual loss function to reduce artifacts introduced by GAN and ensure better visual quality. The experiments on synthetic and real datasets prove that MFAA-GAN is superior to other rain removal models.  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:001120941300003  
Indexed Date:  
2023-12-25  
Institution Address:  
Taiyuan Univ Sci & Technol, Sch Comp Sci & Technol, Taiyuan 030024, Peoples R China  
Institution Address:  
Shanxi Key Lab Adv Control & Equipment Intelligenc, Taiyuan, Peoples R China  
Institution Address:  
North Univ China, Sch Comp Sci & Technol, Taiyuan 030051, Peoples R China  
Institution Address:  
Cent South Univ, Sch Civil Engn, Changsha 410083, Peoples R China  
Institution:  
Institution:  
Shanxi Key Lab Adv Control & Equipment Intelligenc  
Institution:  
Institution:

### 246. Lightweight-Fabric Defect Detection Based on Adaptive Fusion of

摘要: Fabric defect detection is an essential process in the textile industry for ensuring product quality.A algorithm for detecting lightweight fabric defects based on multiscale feature adaptive fusion is proposed to minimize small-target defect detection difficulties in fabric defect detection and the significant differences in the aspect ratios of different defects and high requirements for real-time performance.Based on the YOLOv4 network,lightweight network MobileNetv2 is used as the backbone network to effectively reduce the total number of model parameters and the cost of calculations to satisfy real-time requirements.In the inverse residual structure of MobileNetv2,a new attention mechanism is added.CoordAttention module, which embeds the spatially accurate position information necessary for detecting small-target defects into the channel attention,is used to enhance the ability of the network to focus on small-target features.Second,the Adaptive Spatial Feature Fusion(ASFF)network is used to improve Path Aggregation Network(PANet)to enable the model to obtain the fusion weights of multiscale feature maps through learning,fully utilize shallow and deep features,and further improve the detection accuracy of small-target defects.For different defect types and problems with significant aspect ratio differences and unbalanced categories,the K-means++ algorithm is used to determine the prior frame size,and the Focal Loss function is used to modify the model loss function to reduce the impact of the positive and negative sample imbalance on the detection result. The experimental results show that compared with YOLOv4,the mean Average Precision(mAP)of the proposed algorithm increases by 2.3 percentage points,and the detection speed increases by 12 frame/s. The proposed approach can be effectively applied to fabric defect detection.  
摘要:  
织物瑕疵检测是纺织行业保证产品质量的重要环节,针对织物瑕疵检测中存在小目标瑕疵检测困难、不同种类瑕疵长宽比差异大、对实时性要求高等问题,提出一种新的轻量化织物瑕疵检测算法。以YOLOv4网络为基础,使用轻量化网络MobileNetv2为主干网络,有效减少模型参数总量与运算量,以满足实时性需求。在MobileNetv2的逆残差结构中加入CoordAttention注意力模块,将空间精确位置信息嵌入到通道注意力中,增强网络聚焦小目标特征的能力。使用自适应空间特征融合(ASFF)网络改进路径聚合网络(PANet),使模型通过学习获得多尺度特征图的融合权重,从而充分利用浅层特征与深层特征,提高算法对小目标瑕疵的检测精度。采用K-means++算法确定先验框尺寸,并用Focal Loss函数修改模型损失函数,降低正、负样本不平衡对检测结果的影响,解决不同种类瑕疵长宽比差异大及类别不平衡的问题。实验结果表明,相较于YOLOv4算法,所提算法的平均精度均值提高了2.3个百分点,检测速度提升了12 frame/s,能较好地应用于织物瑕疵检测。  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1000-3428  
入藏号:  
CSCD:7356512  
Indexed Date:  
2023-03-31  
Institution Address:  
江南大学人工智能与计算机学院, 无锡, 江苏 214122, 中国  
Institution Address:  
School of Artificial Intelligence and Computer, Jiangnan University, Wuxi, Jiangsu 214122, China  
Institution:  
江南大学人工智能与计算机学院  
Institution:  
School of Artificial Intelligence and Computer, Jiangnan University

### 247. Enhanced lightweight deep network for efficient livestock detection in

摘要: There are problems in the special pastoral environment, including large changes in target size and serious interference from light and environmental factors. To solve the above problems, an enhanced YOLOv4-tiny target detection network is proposed in this study. This network first solves the problem of livestock size fluctuation in pastoral areas, uses a pyramid network with multiscale feature fusion, and considers shallow local detail features and deep semantic information. Subsequently, a novel compound multichannel attention mechanism is proposed to increase the accuracy of the target detection network for the pastoral environment. The problem of poor accuracy of target detection network is solved. The algorithm is ported to Jetson AGX embedded platform for validation to examine the real-time performance of the algorithm. As revealed by the experimental results, enhanced YOLOv4-tiny achieves 89.77% detection accuracy and 30 frames/second detection speed, which increases the average detection accuracy by 11.67% compared with the conventional YOLOv4-tiny while maintaining almost the same detection rate.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1729-8814  
入藏号:  
WOS:001139109500001  
Indexed Date:  
2024-01-15  
Institution Address:  
Inner Mongolia Univ Technol, Inst Elect Power, Hohhot, Peoples R China  
Institution Address:  
Ordos Vocat Coll, Ordos, Peoples R China  
Institution Address:  
Intelligent Energy Technol & Equipment Engn Res Ct, Hohhot, Peoples R China  
Institution Address:  
Minist Water Resources, Pastoral Water Conservancy Res Inst, Hohhot, Peoples R China  
Institution Address:  
Inner Mongolia Univ Technol, Hohhot 010051, Inner Mongolia, Peoples R China  
Institution:  
Institution:  
Ordos Vocat Coll  
Institution:  
Intelligent Energy Technol & Equipment Engn Res Ct  
Institution:  
Minist Water Resources  
Institution:

### 248. MMFF-NET: Multi-layer and multi-scale feature fusion network for

摘要: Most existing infrared image enhancement algorithms focus on detail and contrast enhancement of ordinary infrared images, and when applied to low-light infrared images, detail and target texture are often severely lost. The reason is that most algorithms process images in a single scale and have difficulty coping with the degradation of image features while enhancing brightness. To solve this problem, we propose a multi-layer and multi-scale feature fusion network (MMFF-Net). It can improve the brightness of low-light infrared images in the absence of normal-light reference samples and keep the image details consistent with the source image. In this paper, features at different layers of the image are extracted using an adaptively modified deep network. A multi-scale adaptive feature fusion module (MAFFM) is designed to preserve and fuse multi-scale information from different convolutional layer features. The fusion features are passed to the iterative function as pixel-wise parameters for image brightness enhancement. We also propose the local feature fusion module (LFFM), which reconstructs images after fusing multiple features, including brightness enhancement images and source images. Finally, in order to implement the training of the whole network, a set of loss functions is carefully designed in this paper. After extensive experiments, it is shown that the algorithm in this paper can effectively enhance low-light infrared images and perform well in subjective visual tests and quantitative tests compared to existing methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
4  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001091311900006  
Indexed Date:  
2023-11-15  
Institution Address:  
Chongqing Univ Technol, Sch Mech Engn, Chongqing 400054, Peoples R China  
Institution Address:  
Chongqing Univ Technol, Engn Res Ctr Mech Testing Technol & Equipment, Minist Educ, Chongqing 400054, Peoples R China  
Institution Address:  
Chongqing Univ Technol, Chongqing Key Lab Time Grating Sensing & Adv Testi, Chongqing 400054, Peoples R China  
Institution:  
Institution:  
Institution:

### 249. Improving the Detection and Positioning of Camouflaged Objects in YOLOv8

摘要: Camouflaged objects can be perfectly hidden in the surrounding environment by designing their texture and color. Existing object detection models have high false-negative rates and inaccurate localization for camouflaged objects. To resolve this, we improved the YOLOv8 algorithm based on feature enhancement. In the feature extraction stage, an edge enhancement module was built to enhance the edge feature. In the feature fusion stage, multiple asymmetric convolution branches were introduced to obtain larger receptive fields and achieve multi-scale feature fusion. In the post-processing stage, the existing non-maximum suppression algorithm was improved to address the issue of missed detection caused by overlapping boxes. Additionally, a shape-enhanced data augmentation method was designed to enhance the model's shape perception of camouflaged objects. Experimental evaluations were carried out on camouflaged object datasets, including COD and CAMO, which are publicly accessible. The improved method exhibits enhancements in detection performance by 8.3% and 9.1%, respectively, compared to the YOLOv8 model.  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
3  
eISSN:  
2079-9292  
入藏号:  
WOS:001096007500001  
Indexed Date:  
2023-11-27  
Institution Address:  
Peoples Liberat Army Engn Univ, Inst Command & Control Engn, Nanjing 210007, Peoples R China  
Institution Address:  
Army Artillery & Def Acad PLA, Nanjing 210007, Peoples R China  
Institution Address:  
China Elect Technol Grp Corp, Res Inst 28, Nanjing 210007, Peoples R China  
Institution:  
Peoples Liberat Army Engn Univ  
Institution:  
Army Artillery & Def Acad PLA  
Institution:

### 250. A Non-Local Attention Feature Fusion Network for Multiscale Object

摘要: Feature pyramid networks (FPN) provide typical architectures for building learning networks with advanced semantic features, which are essential for object recognition at different scales. However, FPN have severe shortcomings in the feature extraction and fusion stages, such as making the extracted features lack of contextual and deep semantic information. In this work, we propose a non-local channel and spatial attention feature pyramid network (NCS-FPN) to improve multi-scale learning. In the feature extraction process, contextual semantic information from different scales is collected through non-local attention networks. In the feature fusion phase, deeper feature information from both spatial and context-aware sources is aggregated to enhance multi-scale feature extraction. Extensive experiments are carried out on two public datasets, MS COCO and PASCAL VOC, are carried out and the results demonstrate that NCS-FPN achieves better performance than several SOTA methods in most cases.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
eISSN:  
2469-7281  
入藏号:  
WOS:000899250200009  
Indexed Date:  
2023-01-12  
Institution Address:  
Chinese Acad Sci, Inst Automat, Beijing 100000, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Artificial Intelligence, Beijing 101408, Peoples R China  
Institution Address:  
Chinese Acad Sci, Guangdong Engn Res Ctr 3D Printing & Intelligent M, Cloud Comp Ctr, Dongguan 523808, Peoples R China  
Institution:  
Institution:  
Institution:

### 251. Cross-Attention Guided Group Aggregation Network for Cropland Change

摘要: Cropland resources are essential for the provision of food production, which is one of the most fundamental needs of human life. Change detection (CD) technology enables the dynamic monitoring of high-resolution cropland resource images acquired through remote sensing satellite sensors. However, current CD methods are not capable of extracting meaningful change information from dense and continuously distributed cropland. In addition, the common feature fusion processing often results in information redundancy and the loss of key features. Therefore, we propose a cross-attention guided group aggregation network (CAGNet) to achieve effective cropland CD. Specifically, we adopt a cross-attention (CA) module to enhance the capability of extracting and characterizing the features of the changed region, reducing the influence of noise and pseudo-change on CD. To alleviate the loss of key information during the multiscale feature fusion process and thus improve the CD performance, we design a group aggregation (GA) module that gradually groups and aggregates the bitemporal features from coarse to fine. Finally, we use a fully convolutional network to obtain the detailed CD results. Furthermore, we demonstrate the effectiveness of knowledge transfer in the field of CD. It allows the models to obtain the underlying mechanisms and characterization capabilities of changed features on the building CD dataset in advance, which significantly improves the performance of various methods on the cropland CD dataset. The experimental results show that CAGNet's quantitative metrics results on the cropland dataset (CL-CD) outperform the other ten benchmarked methods, achieving an F1-score of 79.53%.  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
19  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
20  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001014626700125  
Indexed Date:  
2023-08-16  
Institution Address:  
Hubei Univ Technol, Sch Comp Sci, Wuhan 430068, Peoples R China  
Institution Address:  
Naval Engn Univ, Sch Weapon Engn, Wuhan 430030, Peoples R China  
Institution Address:  
Wuchang Shouyi Univ, Sch Informat Sci & Engn, Wuhan 430032, Peoples R China  
Institution:  
Institution:  
Naval Engn Univ  
Institution:

### 252. Vehicle detection method based on adaptive multi-scale feature fusion

摘要: To improve the vehicle detection accuracy and solve the problem that small vehicles are difficult to detect, an adaptive multi-scale feature fusion network (AMFFN) is proposed to deal with the multi-scale problem, and better performance is achieved after applying it to you only look once (YOLO) v4. To improve the representation capability of features, spatial pyramid pooling modules were employed on each feature map. The proposed AMFFN fuses features of multiple scales across layers and assigns learnable weights to layers of different scales. To achieve detailed information better, we select dynamic rectified linear unit as the activation function, which can change dynamically with the input. AMFFN can be treated as a reusable module to obtain more refined features by repeatedly fusing features. To avoid the huge amount of parameters caused by the complex network, depthwise separable convolution is used to replace the normal convolution and increase the speed of detection. Experimental results show that the proposed method has higher detection accuracy and faster detection speed, and the performance is better than that of YOLO v5, which is the latest version of the YOLO series of algorithms.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:000848751400026  
Indexed Date:  
2022-09-09  
Institution Address:  
Jilin Univ, Coll Comp Sci & Technol, Changchun, Peoples R China  
Institution Address:  
Jilin Univ, Coll Software, Changchun, Peoples R China  
Institution:  
Institution:

### 253. A Multi-Scale Feature Fusion Based Lightweight Vehicle Target Detection

摘要: Vehicle detection with optical remote sensing images has become widely applied in recent years. However, the following challenges have remained unsolved during remote sensing vehicle target detection. These challenges include the dense and arbitrary angles at which vehicles are distributed and which make it difficult to detect them; the extensive model parameter (Param) that blocks real-time detection; the large differences between larger vehicles in terms of their features, which lead to a reduced detection precision; and the way in which the distribution in vehicle datasets is unbalanced and thus not conducive to training. First, this paper constructs a small dataset of vehicles, MiVehicle. This dataset includes 3000 corresponding infrared and visible image pairs, offering a more balanced distribution. In the infrared part of the dataset, the proportions of different vehicle types are as follows: cars, 48%; buses, 19%; trucks, 15%; freight, cars 10%; and vans, 8%. Second, we choose the rotated box mechanism for detection with the model and we build a new vehicle detector, ML-Det, with a novel multi-scale feature fusion triple cross-criss FPN (TCFPN), which can effectively capture the vehicle features in three different positions with an mAP improvement of 1.97%. Moreover, we propose LKC-INVO, which allows involution to couple the structure of multiple large kernel convolutions, resulting in an mAP increase of 2.86%. We also introduce a novel C2F\_ContextGuided module with global context perception, which enhances the perception ability of the model in the global scope and minimizes model Params. Eventually, we propose an assemble-disperse attention module to aggregate local features so as to improve the performance. Overall, ML-Det achieved a 3.22% improvement in accuracy while keeping Params almost unchanged. In the self-built small MiVehicle dataset, we achieved 70.44% on visible images and 79.12% on infrared images with 20.1 GFLOPS, 78.8 FPS, and 7.91 M. Additionally, we trained and tested our model on the following public datasets: UAS-AOD and DOTA. ML-Det was found to be ahead of many other advanced target detection algorithms.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
2072-4292  
入藏号:  
WOS:001332816700001  
Indexed Date:  
2024-10-21  
Institution Address:  
Chinese Acad Sci, Changchun Inst Opt Fine Mech & Phys, Changchun 130033, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Optoelect, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 254. MPS-FFA: A multiplane and multiscale feature fusion attention network

摘要: Structural magnetic resonance imaging (sMRI) is a popular technique that is widely applied in Alzheimer's disease (AD) diagnosis. However, only a few structural atrophy areas in sMRI scans are highly associated with AD. The degree of atrophy in patients' brain tissues and the distribution of lesion areas differ among patients. Therefore, a key challenge in sMRI-based AD diagnosis is identifying discriminating atrophy features. Hence, we propose a multiplane and multiscale feature-level fusion attention (MPS-FFA) model. The model has three components, (1) A feature encoder uses a multiscale feature extractor with hybrid attention layers to simultaneously capture and fuse multiple pathological features in the sagittal, coronal, and axial planes. (2) A global attention classifier combines clinical scores and two global attention layers to evaluate the feature impact scores and balance the relative contributions of different feature blocks. (3) A feature similarity discriminator minimizes the feature similarities among heterogeneous labels to enhance the ability of the network to discriminate atrophy features. The MPS-FFA model provides improved interpretability for identifying discriminating features using feature visualization. The experimental results on the baseline sMRI scans from two databases confirm the effectiveness (e.g., accuracy and generalizability) of our method in locating pathological locations. The source code is available at https://github.com/LiuFei-AHU/MPSFFA.  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:000959334400001  
PubMed ID:  
36958239  
Indexed Date:  
2023-04-22  
Institution Address:  
Anhui Univ, Anhui Prov Int Joint Res Ctr Adv Technol Med Imag, Hefei, Peoples R China  
Institution Address:  
Anhui Univ, Sch Comp Sci & Technol, Hefei, Peoples R China  
Institution Address:  
Monash Univ Malaysia, Sch Engn, Kuala Lumpur, Malaysia  
Institution:  
Institution:  
Institution:

### 255. Infrared Small Target Detection Method Based on Multi-Scale Feature

摘要: Infrared small target detection is widely used in aerial target detection and tracking systems owing to its long detection range and strong anti-jamming ability.Aiming at to overcome the shortcomings of the current infrared small target detection algorithm,such as a low precision rate and high false alarm rate when dealing with complex backgrounds,we propose an end-to-end infrared small target detection model(called MFSSD) based on multi-scale feature fusion.Considering the traits of the targets,we propose a feature fusion module using a refinement and fusion feature map method and improve the correlation of different channels through the SP module.The experimental results of three different sequences of infrared image detection show that the average detection accuracy of the MFSSD algorithm for infrared small target detection was as high as 87.8%.Compared with those of the traditional multi-scale target detection algorithm,both the precision rate and recall rate have been significantly improved.  
摘要:  
红外小目标检测因其探测距离远、抗干扰能力强等特点,在空中目标探测与跟踪系统中得到了广泛的应用。针对目前红外小目标检测算法在复杂背景下检测准确率低、虚警率高等缺点。提出了一种基于多尺度特征融合的端到端红外小目标检测模型(multi-scale feature fusion single shot multibox detecto,MFSSD)。考虑到红外小目标的特点,通过细化和融合特征图的方法提出了一种特征融合模块,通过SP模块提高特征图不同通道的相关性,3种不同序列红外图像的实验结果表明,该算法在红外小目标检测中的平均检测精度高达87.8%。与传统的多尺度目标检测算法相比,准确率和召回率都有显著提高。  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1001-8891  
入藏号:  
CSCD:7011152  
Indexed Date:  
2021-10-09  
Institution Address:  
河南师范大学电子与电气工程学院;;中国科学院界面物理技术重点实验室, ;;中国科学院界面物理技术重点实验室, 新乡;;, 河南;;上海 453007;;201800, 中国  
Institution Address:  
河南师范大学电子与电气工程学院, 新乡, 河南 453007, 中国  
Institution Address:  
河南师范大学电子与电气工程学院;;河南省光电传感集成应用重点实验室, ;;河南省光电传感集成应用重点实验室, 新乡;;新乡, 河南;;河南 453007;;453007, 中国  
Institution Address:  
河南省光电传感集成应用重点实验室, 河南省光电传感集成应用重点实验室, 新乡, 河南 453007, 中国  
Institution Address:  
中国科学院界面物理技术重点实验室, 中国科学院界面物理技术重点实验室, 上海 201800, 中国  
Institution Address:  
College of Electronic and Electrical Engineering,Henan Normal University;;Key laboratory of interfacial Physics Technology project,Chinese Academy of Sciences, ;;Key laboratory of interfacial Physics Technology project,Chinese Academy of Sciences, Xinxiang;;, ;;Shanghai 453007;;201800  
Institution Address:  
College of Electronic and Electrical Engineering,Henan Normal University, Xinxiang, Henan 453007, China  
Institution Address:  
College of Electronic and Electrical Engineering,Henan Normal University;;Henan Key Laboratory of Optoelectronic Sensing Integrated Application, ;;Henan Key Laboratory of Optoelectronic Sensing Integrated Application, Xinxiang;;Xinxiang, ;; 453007;;453007  
Institution Address:  
Henan Key Laboratory of Optoelectronic Sensing Integrated Application, Henan Key Laboratory of Optoelectronic Sensing Integrated Application, Xinxiang, Henan 453007, China  
Institution Address:  
Key laboratory of interfacial Physics Technology project,Chinese Academy of Sciences, Key laboratory of interfacial Physics Technology project,Chinese Academy of Sciences, Shanghai 201800, China  
Institution:  
河南师范大学电子与电气工程学院;;中国科学院界面物理技术重点实验室  
Institution:  
河南师范大学电子与电气工程学院  
Institution:  
河南师范大学电子与电气工程学院;;河南省光电传感集成应用重点实验室  
Institution:  
河南省光电传感集成应用重点实验室  
Institution:  
中国科学院界面物理技术重点实验室  
Institution:  
College of Electronic and Electrical Engineering,Henan Normal University;;Key laboratory of interfacial Physics Technology project,Chinese Academy of Sciences  
Institution:  
College of Electronic and Electrical Engineering,Henan Normal University  
Institution:  
College of Electronic and Electrical Engineering,Henan Normal University;;Henan Key Laboratory of Optoelectronic Sensing Integrated Application  
Institution:  
Henan Key Laboratory of Optoelectronic Sensing Integrated Application  
Institution:  
Key laboratory of interfacial Physics Technology project,Chinese Academy of Sciences

### 256. Facial Expression Editing Technology with Fused Feature Coding

摘要: In order to solve the problems that the current continuous facial expression generation model is easy to produce artifacts in the expression-intensive areas and the expression control ability is weak, the GANimation model is improved for increasing the accuracy of the AU control of the expression muscle motor unit. A multi dimension feature fusion (MFF) module is introduced between the encoding and decoding feature layers of the generator, and the obtained fusion features are used for image decoding in a long-hop connection. A layer of inverse convolution is added to the decoding part of the generator to facilitate the addition of the MFF module to be more efficient and reasonable. Comparing experiments with the original network on the self-made data set, the accuracy of expression synthesis and the quality of the generated images of the improved model have been increased by 1.28 and 2.52 respectively, which verifies that the improved algorithm has better performance in facial expression editing when the image is not blurred and artifacts exist.  
摘要:  
为解决当前连续面部表情生成模型易在表情密集区域产生伪影、表情控制能力较弱等问题,该文对GANimation模型进行了研究改进,提高对表情肌肉运动单元AU控制的准确度。在生成器的编码和解码特征层之间引入多尺度特征融合(MFF)模块,以长跳跃连接的方式将得到的融合特征用于图像解码。在生成器的解码部分中加入一层逆卷积,便于MFF模块添加,更加高效合理。在自制的数据集上与原网络进行对比实验,表情合成的准确度和生成的图像质量分别提高了1.28和2.52,验证了该算法在生成图像没有模糊和伪影存在的情况下,面部表情编辑能力得到加强。  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1001-0548  
入藏号:  
CSCD:7072821  
Indexed Date:  
2021-12-31  
Institution Address:  
沈阳理工大学自动化与电气工程学院, 沈阳, 辽宁 110159, 中国  
Institution Address:  
School of Automation and Electrical Engineering,Shenyang Ligong University, Shenyang, Liaoning 110159, China  
Institution:  
沈阳理工大学自动化与电气工程学院  
Institution:  
School of Automation and Electrical Engineering,Shenyang Ligong University

### 257. Enhancing photovoltaic module cells defect detection by lightweight

摘要: The defect detection of photovoltaic module cells is very important in the process of installation, operation and maintenance of photovoltaic system. It directly affects the efficiency of photovoltaic power generation and the long-term stability of the system. The traditional manual detection method is low efficiency, high cost, and vulnerable to human factors. In contrast, computer vision method can realize automatic, efficient and high-precision defect detection. It significantly improves the accuracy and efficiency of detection. In this paper, an enhanced photovoltaic module cell defect detection algorithm named LFPD-YOLO is proposed by lightweight feature extraction and multi-scale fusion network modules. Firstly, a lightweight feature extraction network module is constructed to reduce the number of parameters and calculations. Secondly, a multi-scale feature fusion network module is proposed. The detection ability of small targets is improved. We use mixed local channel attention mechanism to enhance the ability to capture features. We use knowledge distillation to balance the efficiency and accuracy of the network. Given the experimental results, compared with the original network, our algorithm reduces the number of model parameters by 37% and increases the map50 by 0.6%. LFPD-YOLO better improves the accuracy and parametric count of the network. It provides an effective solution for PV module cell defect detection.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1543-5075  
eISSN:  
1543-5083  
入藏号:  
WOS:001435687600001  
Indexed Date:  
2025-03-08  
Institution Address:  
Tianjin Chengjian Univ, Sch Control & Mech, Tianjin 300384, Peoples R China  
Institution Address:  
Harbin Inst Technol, Sch Instrument Sci & Engn, Harbin, Peoples R China  
Institution Address:  
Power Construct Corp China, STECOL Corp, Tianjin, Peoples R China  
Institution:  
Institution:  
Institution:  
Power Construct Corp China

### 258. Multiscale Fusion Crowd Counting Algorithm Based on Attention Mechanism

摘要: A multiscale fusion crowd counting algorithm based on attention mechanism is proposed to addresses the issues of large head scale changes and high background noise in crowd counting images,fully aggregating multiscale information to effectively distinguish background noise.Atrous spatial pyramid pooling based on residual connection method is constructed to capture multiscale head target features while incorporating spatial details from shallow feature maps through residual structures and multiple dilated convolutions with different expansion rates,thereby improving the quality of feature maps.A cross-layer multiscale feature fusion module is built to integrate edge details and contextual semantic information of different sizes of shallow and deep branches.In addition,a feature fusion module based on multi-branch is designed to integrate multiscale information of different receptive field sizes,thereby alleviating the problem of large-scale head scale changes.A channel and spatial attention mechanism module is further constructed based on the matrix similarity operation to extract pixel level feature weights,enhance the network's discriminative ability for background and head targets,and adaptively correct position information.The experimental results show that compared to the optimal values of the 11 comparison algorithms,the proposed algorithm reduces the Mean Absolute Error(MAE)and Root Mean Square Error(RMSE)indicators by 1.4% and 4.2% on the SHA dataset,and reduced by 4.9% and 1.8% on the UCF\_CC\_50 dataset,the proposed algorithm can accurately predict the distribution status,estimate the number of people,and generate high-quality population density maps.  
摘要:  
针对人群计数图像人头尺度变化大、背景噪声高等问题,提出一种基于注意力机制的多尺度融合人群计数算法,以充分聚合多尺度信息,并有效区分背景噪声。构建基于残差连接的空洞空间金字塔池化,通过残差结构以及多个不同扩张率的空洞卷积在捕获多尺度头部目标特征的同时融入浅层特征图的空间细节信息,提高特征图质量;构建跨层多尺度特征融合模块,融合浅层和深层分支不同大小的边缘细节信息和上下文语义信息,并设计基于多分支的特征融合模块,融合不同感受野大小的多尺度信息以缓解大规模人头尺度变化的问题;构建基于矩阵相似运算的通道和空间注意力机制模块提取像素级特征权重,加强网络对于背景和人头目标的判别能力,自适应矫正位置信息。实验结果表明,相比11种对比算法的最优值,所提算法在SHA数据集上的平均绝对误差和均方根误差指标降低1.4%、4.2%,在UCF\_CC\_50数据集上降低4.9%、1.8%,能够精确地预测人群分布状态和估计人群数量,生成高质量的人群密度图。  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1000-3428  
入藏号:  
CSCD:7670028  
Indexed Date:  
2024-05-24  
Institution Address:  
太原科技大学电子信息工程学院;;太原科技大学, ;;先进控制与装备智能化山西省重点实验室, 太原;;太原, 山西;;山西 030024;;030024, 中国  
Institution Address:  
School of Electronic Information Engineering,Taiyuan University of Science and Technology;;Taiyuan University of Science and Technology, ;;Shanxi Key Laboratory of Advanced Control and Equipment Intelligence, Taiyuan;;Taiyuan, Shanxi;;Shanxi 030024;;030024  
Institution:  
太原科技大学电子信息工程学院;;太原科技大学  
Institution:  
School of Electronic Information Engineering,Taiyuan University of Science and Technology;;Taiyuan University of Science and Technology

### 259. CMFCANet: A Wind Turbine Blade Damage Recognition Network Based on

摘要: Drone inspections for detecting wind turbine blade (WTB) defects are efficient but necessitate advanced feature recognition methods for aerial imagery. This article introduces CMFCANet, a novel framework designed to accurately characterize WTB defect features through multifeature fusion. The RGB-HSI color space fusion module facilitates the integration of color information at the feature level, significantly enhancing the extraction of color and shadow details in defect images. Additionally, the proposed network employs a feature pyramid for multiscale feature fusion, thereby improving its adaptability to defects of varying sizes. The channel attention module (CAM) dynamically assigns weights to individual feature channels, emphasizing the most salient features that influence model decisions. Furthermore, class activation maps are utilized to analyze the model's feature representation, enhancing interpretability. To validate the effectiveness of the proposed method, a comprehensive dataset comprising six distinct types of WTB defects, ranging from minor to severe damage, has been constructed. Experimental results indicate that CMFCANet outperforms existing state-of-the-art (SOTA) models, achieving an overall accuracy of 84.211%. Moreover, the proposed model demonstrates balanced performance, with precision, recall, $F1$ score, and kappa metrics that compare favorably to other methods.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001397823700027  
Indexed Date:  
2025-01-22  
Institution Address:  
Univ Shanghai Sci & Technol, Sch Mech Engn, Shanghai 200093, Peoples R China  
Institution:

### 260. MFCANet: A road scene segmentation network based on Multi-Scale feature

摘要: Road scene segmentation is the basic task of autonomous driving. Recent representative scene segmentation methods adopt the full convolutional network based on the encoder -decoder. However, the framework can cause the loss of image fine-grained information in the process of down -sampling, feature extraction and feature fusion, resulting in blurred boundary details and chaotic segmentation effect. In this work, a road scene segmentation network based on multi -scale feature fusion and context information aggregation is proposed, in which context information is used to guide feature fusion and enhance semantic feature extraction. Three plug -and -play modules are designed to extract multi -scale features with strong semantic information from high-level features, which compensate for the loss of spatial information in the upper sampling stage, and capture the information dependence among pixels to improve pixel -by -pixel segmentation. Experimental results on Camvid and Cityscapes show that the proposed multi -scale feature fusion and context information aggregation network (MFCANet) can achieve satisfactory performance compared with the state-of-the-art segmentation methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1047-3203  
eISSN:  
1095-9076  
入藏号:  
WOS:001170574700001  
Indexed Date:  
2024-03-11  
Institution Address:  
Qingdao Univ Sci & Technol, Coll Electromech Engn, Qingdao 266061, Peoples R China  
Institution Address:  
Qingdao Univ Sci & Technol, Sch Math & Phys, Qingdao 266061, Peoples R China  
Institution Address:  
Qingdao Univ, Coll Comp Sci & Technol, Qingdao 266000, Peoples R China  
Institution:  
Institution:  
Institution:

### 261. An Improved Pedestrian Detection Model Based on YOLOv8 for Dense Scenes

摘要: In dense scenes, pedestrians often exhibit a variety of symmetrical features, such as symmetry in body contour, posture, clothing, and appearance. However, pedestrian detection poses challenges due to the mutual occlusion of pedestrians and the small scale of distant pedestrians in the image. To address these challenges, we propose a pedestrian detection algorithm tailored for dense scenarios called YOLO-RAD. In this algorithm, we integrate the concept of receiving field attention (RFA) into the Conv and C2f modules to enhance the feature extraction capability of the network. A self-designed four-layer adaptive spatial feature fusion (ASFF) module is introduced, and shallow pedestrian feature information is added to enhance the multi-scale feature fusion capability. Finally, we introduce a small-target dynamic head structure (DyHead-S) to enhance the capability of detecting small-scale pedestrians. Experimental results on WiderPerson and CrowdHuman, two challenging dense pedestrian datasets, show that compared with YOLOv8n, our YOLO-RAD algorithm has achieved significant improvement in detection performance, and the detection performance of mAP@0.5 has increased by 2.5% and 6%, respectively. The detection performance of mAP@0.5:0.95 was improved by 2.7% and 6.8%, respectively. Therefore, the algorithm can effectively improve the performance of pedestrian detection in dense scenes.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
2073-8994  
入藏号:  
WOS:001256047600001  
Indexed Date:  
2024-07-04  
Institution Address:  
Changchun Univ Technol, Sch Comp Sci & Engn, Changchun 130012, Peoples R China  
Institution:

### 262. SFFNet: Shallow Feature Fusion Network Based on Detection Framework for

摘要: Infrared small target detection (IRSTD) is the process of recognizing and distinguishing small targets from infrared images that are obstructed by crowded backgrounds. This technique is used in various areas, including ground monitoring, flight navigation, and so on. However, due to complex backgrounds and the loss of information in deep networks, infrared small target detection remains a difficult undertaking. To solve the above problems, we present a shallow feature fusion network (SFFNet) based on detection framework. Specifically, we design the shallow-layer-guided feature enhancement (SLGFE) module, which guides multi-scale feature fusion with shallow layer information, effectively mitigating the loss of information in deep networks. Then, we design the visual-Mamba-based global information extension (VMamba-GIE) module, which leverages a multi-branch structure combining the capability of convolutional layers to extract features in local space with the advantages of state space models in the exploration of long-distance information. The design significantly extends the network's capacity to acquire global contextual information, enhancing its capability to handle complex backgrounds. And through the effective fusion of the SLGFE and VMamba-GIE modules, the exorbitant computation brought by the SLGFE module is substantially reduced. The experimental results on two publicly available infrared small target datasets demonstrate that the SFFNet surpasses other state-of-the-art algorithms.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:001366534000001  
Indexed Date:  
2024-12-07  
Institution Address:  
Natl Lab Adapt Opt, Chengdu 610209, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Opt & Elect, Chengdu 610209, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 101408, Peoples R China  
Institution:  
Natl Lab Adapt Opt  
Institution:  
Institution:

### 263. 3D liver image segmentation method based on multi-scale feature fusion

摘要: Due to the high similarity of gray values among liver and adjacent organs in Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) images,a 3D liver image segmentation method based on multi-scale feature fusion and grid attention mechanism,namely MAGNet (Multi-scale feature fusion And Grid attention mechanism Network),was proposed to segment liver automatically and accurately.Firstly,high-level features and low-level features were connected by the attention-guided concatenation module to extract important context information,and the grid attention mechanism was introduced in the attention-guided concatenation module to focus on the segmentation region of interest.Then,the multiscale feature fusion module was formed by the layered connection in a single feature map according to the number of channels,and this module was used to replace the basic convolutional block to obtain multi-scale semantic information.Finally,the deep supervision mechanism was utilized to solve the problems of vanishing gradient,exploding gradient and slow convergence.Experimental results show that on 3DIRCADb dataset,compared with the U3-Net+DC method,MAGNet improves the Dice Similarity Coefficient (DSC) metric by 0.10 percentage points and reduces the Relative Volume Difference (RVD) metric by 1.97 percentage points;on Sliver07 dataset,compared with the CANet method,MAGNet improves the DSC metrics by 0.30 percentage points,reduces Volumetric Overlap Error (VOE) metrics by 0.68 percentage points,and reduces the Average Symmetric Surface Distance (ASD) and Root Mean Square Symmetric Surface Distance(RMSD) metrics 0.03 mm and 0.22 mm respectively;on the liver MRI dataset of a hospital,MAGNet also has good results on all metrics.Besides,MAGNet was applied to a mixed dataset of 3DIRCADb dataset and the hospital liver MRI dataset above,and a competitive segmentation result was also achieved.  
摘要:  
在计算机断层扫描(CT)和磁共振成像(MRI)的影像中肝脏与邻近脏器的灰度值相似性都比较高,为自动精确地分割肝脏,提出一种基于多尺度特征融合和网格注意力机制的三维肝脏影像分割方法MAGNet (Multi-scale feature fusion And Grid attention mechanism Network)。首先,通过注意力引导连接模块来连接高层特征和低层特征以提取出重要的上下文信息,并且在注意力引导连接模块中引入网格注意力机制来关注感兴趣的分割区域;然后,通过在单个特征图中按通道数进行分层连接形成多尺度特征融合模块,并用该模块替换基础卷积块以获取多尺度语义信息;最后,利用深度监督机制解决梯度消失、梯度爆炸和收敛过慢等问题。实验结果表明:在3DIRCADb数据集上,与U3-Net+DC方法相比,MAGNet在Dice相似系数(DSC)指标上提升了0.10个百分点,在相对体积差(RVD)指标上降低了1.97个百分点;在Sliver07数据集上,与CANet方法相比,MAGNet在DSC指标上提升了0.30个百分点,在体素重叠误差(VOE)指标上降低了0.68个百分点,在平均对称表面距离(ASD)和对称位置表面距离的均方根(RMSD)指标上分别降低了0.03 mm和0.22 mm;在某医院肝脏MRI数据集上,MAGNet在所有指标上也均具有良好的结果。另外,将MAGNet应用于3DIRCADb数据集和某医院肝脏MRI数据集进行混合形成的数据集,也取得了非常有竞争力的分割效果。  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1001-9081  
入藏号:  
CSCD:7529978  
Indexed Date:  
2023-11-24  
Institution Address:  
武汉科技大学计算机科学与技术学院;;武汉科技大学大数据科学与工程研究院;;武汉科技大学, ;;;;智能信息处理与实时工业系统湖北省重点实验室, 武汉;;武汉;;武汉, ;;;; 430065;;430065;;430065  
Institution Address:  
武汉科技大学附属天佑医院, 武汉, 湖北 430064, 中国  
Institution Address:  
School of Computer Science and Technology,Wuhan University of Science and Technology;;Institute of Big Data Science and Engineering,Wuhan University of Science and Technology;;Wuhan University of Science and Technology, ;;;;Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan;;Wuhan;;Wuhan, Hubei;;Hubei;;Hubei 430065;;430065;;430065  
Institution Address:  
Tianyou Hospital Affiliated to Wuhan University of Science and Technology, Wuhan, Hubei 430064, China  
Institution:  
武汉科技大学计算机科学与技术学院;;武汉科技大学大数据科学与工程研究院;;武汉科技大学  
Institution:  
武汉科技大学附属天佑医院  
Institution:  
School of Computer Science and Technology,Wuhan University of Science and Technology;;Institute of Big Data Science and Engineering,Wuhan University of Science and Technology;;Wuhan University of Science and Technology  
Institution:  
Tianyou Hospital Affiliated to Wuhan University of Science and Technology

### 264. Scale-Aware Squeeze-and-Excitation for Lightweight Object Detection

摘要: Lightweight object detection can promote intelligent robotics to recognize surroundings objects with limited computational resources, and thus receives increasing attention in robotics communities. Recently, high-resolution networks (HRNets) can learn high-resolution representation and it obtains excellent performance as the backbones of current cutting-edge object detectors. However, two crucial issues remain with regard to applying HRNet-based detectors to mobile devices-insufficient local feature interactions and multiscale feature fusion. In this work, we propose a scale-aware squeeze-and-excitation (SASE) module that utilizes SE operations to fully explore feature interactions without increasing network complexity; this is followed by a scale-aware attention (SAA) mechanism, which adaptively fuses multiscale features by estimating the importance of each scale. The SASE module can serve as the basic block for the HRNet, which facilitates the use of HRNet as a backbone for lightweight object detection. Extensive experiments conducted on Microsoft COCO and Pascal VOC demonstrate that the proposed method has a good tradeoff between accuracy and model complexity. With similar numbers of parameters and calculations, the mean average precision (mAP) achieved on the COCO dataset is improved by 3.7% over that of Lite-HRNet.  
Web of Science 核心合集中的 "被引频次":  
9  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
9  
ISSN:  
2377-3766  
入藏号:  
WOS:000890837400007  
Indexed Date:  
2022-12-17  
Institution Address:  
Guangdong Univ Technol, Sch Informat Engn, Guangzhou 510006, Peoples R China  
Institution Address:  
South China Univ Technol, Sch Mech & Automot Engn, Guangzhou 510641, Peoples R China  
Institution:  
Institution:

### 265. DFNet-Trans: An end-to-end multibranching network for depth estimation

摘要: Transparent objects play a vital role in modern industries and find widespread applications across various engineering scenarios. However, capturing accurate depth maps of transparent objects remains challenging due to their reflective and refractive properties, which pose difficulties for most commercial -grade optical sensors. In this paper, we propose a novel depth estimation method called DFNet-Trans, designed to estimate depth from a noisy RGB-D image input. Initially, a multiscale feature fusion module (FFM) is incorporated into the existing depth estimation network to generate the initial depth map. Subsequently, we enhance the network by adding a confidence branch and a mask branch on the same encoder, enabling improved distortion correction and real scene restoration in the depth estimation. Based on the framework representation, missing depth can be completed. Comprehensive experiments demonstrate that the proposed approach significantly outperforms the current state-of-the-art methods on the recently popular large-scale real dataset TransCG. the proposed approach achieves a remarkable 27.7% reduction in RMSE and a notable 34.6% reduction in REL. The generalization experiment shows that the proposed approach outperforms existing methods when generalized to an unknown real dataset.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1077-3142  
eISSN:  
1090-235X  
入藏号:  
WOS:001166353800001  
Indexed Date:  
2024-03-06  
Institution Address:  
Southwest Jiaotong Univ, Sch Mech Engn, Chengdu 610031, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, Inst Smart City & Intelligent Transportat, Chengdu 610031, Peoples R China  
Institution:  
Institution:

### 266. Image Super-Resolution Based on Residual Attention and Multi-Scale

摘要: At present, deep residual network has been widely used in image super-resolution and proved to be able to achieve good reconstruction results. However, the existing super-resolution algorithms based on deep residual network have the problems of indiscriminately learning feature information of different regions and low utilization rate of feature information, which make them difficult to further improve the reconstruction effect. In view of the above problems, a novel super-resolution reconstruction network based on residual attention and multi-scale feature fusion (RAMF) is proposed in this paper. Firstly, a lightweight multi-scale residual module (LMRM) is proposed in the deep feature extraction stage, by which the multi-scale features are extracted and further cross-connected to enrich the information of different receptive fields. Then, to fully improve the utilization rate of feature information, a dense feature fusion structure is designed to fuse the output feature of each LMRM. Finally, a residual spatial attention module (RSAM) is proposed to specifically learn and better retain high-frequency feature information, so as to improve the reconstruction effect. Experimental tests and comparisons are conducted with the current advanced methods on four baseline databases, and the results demonstrate that the proposed RAMF can achieve better reconstruction effect with fewer parameters, low computational complexity, fast processing speed and high objective evaluation index. Especially, the peak signal-to-noise ratio measured on Urban100 data set increases by 0.13dB on average, and the reconstructed image has better visual effect and richer texture detail features.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001016806000001  
Indexed Date:  
2023-07-07  
Institution Address:  
China Univ Min & Technol, Sch Comp Sci & Technol, Xuzhou 221116, Peoples R China  
Institution Address:  
China Univ Min & Technol, Sch Informat & Control Engn, Xuzhou 221116, Peoples R China  
Institution:  
Institution:

### 267. Multi-Scale Feature Fusion Attention Network for Infrared Small Target

摘要: Compared with other target detection tasks, infrared small target detection has the problem of feature information loss in deep networks due to fewer target pixels and the lack of color and texture features. To address aforementioned issue, a Multi-Scale Feature Fusion Attention Network (MSFFA) is proposed to better utilize shallow edge features and deep semantic features. Its main components contain Convolutional Block Attention Module (CBAM), Multi-Scale Receptive Field Feature Fusion Module (R3FM), and Bidirectional Feature Aggregation Network (BFANet). CBAM is designed to calculate the importance of each feature map and enhance useful features from the channel and spatial dimensions. R3FM is proposed to characterize the global context information of deep layers feature map to enlarge the network's receptive field for small targets detection with a larger range of location information. BFANet is developed to shorten the path of information exchange between different layers and reinforce the utilization of shallow features in the network. Moreover, the K-means clustering algorithm is adopted to optimize the width to height ratio of the bounding anchor, and it can better match the positive samples to improve the training performance. Extensive experiments on public infrared small target detection dataset demonstrate that the proposed method achieves better performance compared to the other state-of-the-art methods.  
会议名称:  
14th International Conference on Graphics and Image Processing (ICGIP)  
会议日期:  
OCT 21-23, 2022  
会议地点:  
ELECTR NETWORK  
会议赞助商:  
Nanjing Univ Sci & Technol, Sch Comp Sci & Engn; Jiangsu Comp Soc; Jiangsu Assoc Artificial Intelligence  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0277-786X  
eISSN:  
1996-756X  
ISBN:  
978-1-5106-6631-3; 978-1-5106-6632-0  
入藏号:  
WOS:001022021800006  
Indexed Date:  
2023-07-27  
Institution Address:  
Zhongyuan Univ Technol, Sch Elect & Informat Engn, Zhengzhou, Peoples R China  
Institution:

### 268. Colorful Image Colorization with Classification and Asymmetric Feature

摘要: An automatic colorization algorithm can convert a grayscale image to a colorful image using regression loss functions or classification loss functions. However, the regression loss function leads to brown results, while the classification loss function leads to the problem of color overflow and the computation of the color categories and balance weights of the ground truth required for the weighted classification loss is too large. In this paper, we propose a new method to compute color categories and balance the weights of color images. In this paper, we propose a new method to compute color categories and balance weights of color images. Furthermore, we propose a U-Net-based colorization network. First, we propose a category conversion module and a category balance module to obtain the color categories and to balance weights, which dramatically reduces the training time. Second, we construct a classification subnetwork to constrain the colorization network with category loss, which improves the colorization accuracy and saturation. Finally, we introduce an asymmetric feature fusion (AFF) module to fuse the multiscale features, which effectively prevents color overflow and improves the colorization effect. The experiments show that our colorization network has peak signal-to-noise ratio (PSNR) and structure similarity index measure (SSIM) metrics of 25.8803 and 0.9368, respectively, for the ImageNet dataset. As compared with existing algorithms, our algorithm produces colorful images with vivid colors, no significant color overflow, and higher saturation.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:000873512300001  
PubMed ID:  
36298360  
Indexed Date:  
2022-11-04  
Institution Address:  
Chinese Acad Sci, Changchun Inst Opt Fine Mech & Phys, Changchun 130033, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 269. YOLO-BGS Optimizes Textile Production Processes: Enhancing YOLOv8n with

摘要: Timely detection of fabric defects is crucial for improving fabric quality and reducing production losses for companies. Traditional methods for detecting fabric defects face several challenges, including low detection efficiency, poor accuracy, and limited types of detectable defects. To address these issues, this paper chose the YOLOv8n model for continuous iteration enhancement in order to improve its detection performance. First, multiscale feature fusion was realized by the Bi-directional Feature Pyramid Network (BiFPN). Second, the Shuffle Attention Mechanism (SA) is introduced to optimize feature classification. Finally, the Global Attention Mechanism (GAM) was used to improve global detection accuracy. Empirical findings demonstrated the improved model's efficacy, attaining a test set mean average precision (mAP) value of 96.6%, which is an improvement of 3.6% compared to the original YOLOv8n. This validates that YOLO-BGS excels in detecting textile defects. It effectively locates these defects, minimizes resource waste, and fosters sustainable production practices.  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
2071-1050  
入藏号:  
WOS:001323601100001  
Indexed Date:  
2024-10-07  
Institution Address:  
Taiyuan Univ Technol, Coll Text Engn, Jinzhong 030600, Peoples R China  
Institution:

### 270. Multiscale Integration Network With Quaternion Convolution for

摘要: In this letter, we proposed a multiscale integration network with quaternion convolution (MQ-Net) for the fusion of low spatial resolution multispectral (LRMS) and panchromatic (PAN) images. In this network, LRMS and PAN images are resampled at different scales and fed into feature fusion modules (FFMs) to merge the spatial and spectral information among them. Then, multiscale feature enhancement modules (MFEMs) are designed to sufficiently learn the spatial and spectral information at different scales. Meanwhile, we employ a quaternion convolution module (QCM) to better capture the dependencies within spectral bands of LRMS images. Then, the quaternion features are introduced into MFEMs for efficient feature enhancement. Finally, all information from different scales is integrated for the reconstruction of high LRMS images. Reduced- and full-resolution experiments are performed on GeoEye-1 and WorldView-2 satellite datasets. Compared to some state-of-the-art pansharpening methods, the proposed MQ-Net obtains better results in terms of qualitative and quantitative evaluations. The code is available at https://github.com/RSMagneto/MQ-Net.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001375803200021  
Indexed Date:  
2024-12-18  
Institution Address:  
Shandong Normal Univ, Sch Informat Sci & Engn, Jinan 250358, Peoples R China  
Institution Address:  
TongJi Univ, Sch Phys Sci & Engn, Shanghai 200092, Peoples R China  
Institution Address:  
Xianyang Normal Univ, Sch Comp Sci, Xianyang 712000, Peoples R China  
Institution:  
Institution:  
Institution:

### 271. Grasp Detection with Hierarchical Multi-Scale Feature Fusion and

摘要: Grasp detection plays a critical role for robot manipulation. Mainstream pixel-wise grasp detection networks with encoder-decoder structure receive much attention due to good accuracy and efficiency. However, they usually transmit the high-level feature in the encoder to the decoder, and low-level features are neglected. It is noted that low-level features contain abundant detail information, and how to fully exploit low-level features remains unsolved. Meanwhile, the channel information in high-level feature is also not well mined. Inevitably, the performance of grasp detection is degraded. To solve these problems, we propose a grasp detection network with hierarchical multi-scale feature fusion and inverted shuffle residual. Both low-level and high-level features in the encoder are firstly fused by the designed skip connections with attention module, and the fused information is then propagated to corresponding layers of the decoder for in-depth feature fusion. Such a hierarchical fusion guarantees the quality of grasp prediction. Furthermore, an inverted shuffle residual module is created, where the high-level feature from encoder is split in channel and the resultant split features are processed in their respective branches. By such differentiation processing, more high-dimensional channel information is kept, which enhances the representation ability of the network. Besides, an information enhancement module is added before the encoder to reinforce input information. The proposed method attains 98.9% and 97.8% in image-wise and object-wise accuracy on the Cornell grasping dataset, respectively, and the experimental results verify the effectiveness of the method.  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1007-0214  
eISSN:  
1878-7606  
入藏号:  
WOS:001074356700020  
Indexed Date:  
2023-10-14  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Multimodal Artificial Intelligence, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Artificial Intelligence, State Key Lab Multimodal Artificial Intelligence, Beijing 100049, Peoples R China  
Institution Address:  
Peking Univ, Coll Engn, Dept Adv Mfg & Robot, Beijing 100871, Peoples R China  
Institution:  
Institution:  
Institution:

### 272. Improved Lightweight Underwater Target Detection Algorithm of YOLOv7

摘要: Aiming at the problems of target false and missing detection caused by limited memory and computing power of underwater equipment and complex underwater environment, a lightweight underwater target detection method YOLOv7-SDBB is proposed. The ShuffleNetv2 lightweight network is introduced on the backbone network of YOLOv7 to reduce the parameter amount and calculation amount of the feature extraction network. The D-ELAN and D-MPConv modules are designed to further realize the network lightweight and improve the model detection speed. Due to the phenomenon of false and missing detection is prone to occur during underwater detection, BiFPN is used to perform multi-scale feature fusion and integrate deep feature information. In view of the problem of feature information loss caused by BiFPN feature fusion, the BiFormer attention mechanism is used to retain key information and improve target detection accuracy. The experimental results show that the accuracy of the improved model on the URPC2020 dataset has increased by 2.7 percentage points, the amount of parameters and calculations have decreased by 20.3% and 41.7% respectively, and the detection speed has increased to 100.9 FPS, realizing a good balance between the speed and accuracy of underwater target detection.  
摘要:  
针对水下设备内存和计算能力有限和水下环境复杂造成的目标错检和漏检问题,提出一种轻量级水下目标检测方法YOLOv7-SDBB。在YOLOv7的骨干网络上引入ShuffleNetv2轻量级网络,降低特征提取网络的参数量和计算量;设计了D-ELAN和D-MPConv模块,在进一步实现网络轻量化的同时提高模型检测速度;由于水下检测过程中容易出现错检、漏检的现象,利用BiFPN(bidirectional feature pyramid network)进行多尺度特征融合,融合深层的特征信息;针对BiFPN特征融合导致的特征信息丢失的问题,采用BiFormer注意力机制保留关键信息,提高目标检测精度。实验结果表明,改进后模型在URPC2020数据集上的精度提高了2.7个百分点,参数量和计算量分别下降了20.3%和41.7%,检测速度提升至100.9 FPS,从而验证了提出的算法在精度和速度之间取得了很好的平衡。  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1002-8331  
入藏号:  
CSCD:7664149  
Indexed Date:  
2024-05-17  
Institution Address:  
西安邮电大学,电子工程学院, 西安, 陕西 710121, 中国  
Institution Address:  
School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Xi'an, Shaanxi 710121, China  
Institution:  
西安邮电大学,电子工程学院  
Institution:  
School of Electronic Engineering, Xi'an University of Posts and Telecommunications

### 273. AEM-YOLOv8s:Small Target Detection Algorithm for UAV Aerial Images

摘要: The AEM-YOLOv8s algorithm is proposed to address issues of low performance, missed detections, occlusions, and high model parameter count in small object detection in current UAV aerial imagery. Within the C2f module, the advantages of AKConv (alterable kernel convolution) and EMA (efficient multi-scale attention) are combined to design the C2f-BE module, which enhances the algorithm's ability to process features while reducing the model parameter count. By introducing a small object detection layer and BiFPN structure, through cross-scale connections and weighted feature fusion, more shallow features are retained, reducing algorithm parameters. The design of a multi-scale feature fusion branch merges shallow features containing more small object information with deeper semantic features, reducing missed detections under occlusion and improving small object detection performance. Experimental results on the VisDrone2019 public dataset demonstrate that the AEM-YOLOv8s algorithm achieves an mAP50 of 50.1% and mAP50:95 of 31.1%, representing respective improvements of 10.8 and 7.6 percentage points over YOLOv8s, while also reducing parameters by 32.2% compared to YOLOv8s.  
摘要:  
针对目前无人机航拍图中的小目标检测性能低、漏检、遮挡以及模型参数量大的问题,提出了AEM-YOLOv8s算法。在C2f模块中结合AKConv(alterable kernel convolution)和EMA(efficient multi-scale attention)的优点,设计了C2f-BE模块,更好地提高了算法处理特征的能力,同时也降低了模型参数量。引入小目标检测层和BiFPN结构,通过跨尺度连接方式和加权特征融合,能够保留更多的浅层特征,并且减少了算法参数量。设计多尺度特征融合分支,将浅层特征与深层特征进行融合,减少了遮挡情况下的漏检,提高了算法对小目标检测性能。在VisDrone2019公开数据集上的实验表明,AEM-YOLOv8s算法的mAP50为50.1%,mAP50:95为31.1%,较YOLOv8s分别提高了10.8和7.6个百分点,同时参数量较YOLOv8s降低了32.2%。  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7804358  
Indexed Date:  
2024-11-23  
Institution Address:  
上海电力大学电子与信息工程学院, 上海 201306, 中国  
Institution Address:  
上海电力大学电子与信息工程学院;;上海电机学院, ;;, ;;, 上海;;上海 201306;;201306, 中国  
Institution Address:  
College of Electronics and Information Engineering,Shanghai University of Electric Power, Shanghai 201306, China  
Institution Address:  
College of Electronics and Information Engineering,Shanghai University of Electric Power;;Shanghai Dianji University, ;;, ;;, Shanghai;;Shanghai 201306;;201306  
Institution:  
上海电力大学电子与信息工程学院  
Institution:  
上海电力大学电子与信息工程学院;;上海电机学院  
Institution:  
College of Electronics and Information Engineering,Shanghai University of Electric Power  
Institution:  
College of Electronics and Information Engineering,Shanghai University of Electric Power;;Shanghai Dianji University

### 274. Object Detection Algorithm in Remote Sensing Images Based on Improved

摘要: Remote sensing target detection is an important aspect in the fields of environmental monitoring and circuit patrol. A remote sensing target detection algorithm based on YOLOX is proposed for the difficulties of remote sensing images with large target scale differences, blurred targets and high background complexity. First, a regional context aggregation module is proposed to expand the perceptual field using the dalited convolutions with different expansion rates to obtain multi-scale contextual information, which is beneficial to the detection of the small targets. Second, the feature fusion module is proposed, and two different scale transformation modules are used to achieve the fusion of features at different scales, fully fusing shallow location information with deep semantic information to improve the detection performance of the network for targets at different scales. Finally, a feature enhancement module is introduced to the multiscale feature fusion network part and combined with the attention mechanism CAS [CA (coordinate attention) with SimAM (simple parameter-free attention module)] to make the network pay more attention to the target information and ignore the interference of complex background, while the shallow feature layer is fused with the deep detection layer for feature fusion to prevent the low detection performance affected by the loss of feature information at the prediction end. The experimental results show that the improved algorithm achieves 73.87% and 96.22% detection accuracy on DIOR and RSOD remote sensing datasets, which is 4.08 and 1.34 percentage points higher than the original YOLOX algorithm, and has superiority in both detection accuracy and detection speed compared with other advanced algorithms.  
摘要:  
遥感目标检测是环境监测、电路巡检等领域中的一个重要的环节。针对遥感图像存在的目标尺度差异大、目标模糊、背景复杂度高等难点,提出了一种基于YOLOX的遥感目标检测算法。首先,提出区域上下文聚合模块,利用不同扩张率的空洞卷积扩大感受野,获取多尺度上下文信息,从而有利于小目标的检测;其次,提出特征融合模块,采用两种不同的尺度变换模块实现对不同尺度特征的融合,从而充分融合浅层位置信息与深层语义信息,提高网络对不同尺度目标的检测性能;最后在多尺度特征融合网络部分引入特征增强模块,并将其与注意力机制CAS[CA(coordinate attention)+SimAM (simple parameter-free attention module)]结合,使网络更加关注目标信息,忽略复杂背景的干扰,同时,将浅层特征层与深层检测层进行特征融合,防止由特征信息丢失造成的预测端检测性能降低。实验结果表明:改进后的算法在DIOR和RSOD遥感数据集上分别取得了73.87%和96.22%的检测精度,与原YOLOX算法相比检测精度分别提高了4.08和1.34个百分点;与其他先进算法相比,改进后的算法在检测精度与检测速度上都具有一定的优越性。  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7783051  
Indexed Date:  
2024-10-18  
Institution Address:  
南京信息工程大学电子与信息工程学院;;南京信息工程大学, ;;江苏省大气环境与装备技术协同创新中心, 南京;;南京, 江苏;;江苏 210044;;210044, 中国  
Institution Address:  
南京信息工程大学电子与信息工程学院, 南京, 江苏 210044, 中国  
Institution Address:  
School of Electronic and Information Engineering,Nanjing University of Information Science and Technology;;Nanjing University of Information Science and Technology, ;;Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, Nanjing;;Nanjing, Jiangsu;;Jiangsu 210044;;210044  
Institution Address:  
School of Electronic and Information Engineering,Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China  
Institution:  
南京信息工程大学电子与信息工程学院;;南京信息工程大学  
Institution:  
南京信息工程大学电子与信息工程学院  
Institution:  
School of Electronic and Information Engineering,Nanjing University of Information Science and Technology;;Nanjing University of Information Science and Technology  
Institution:  
School of Electronic and Information Engineering,Nanjing University of Information Science and Technology

### 275. SiamMFF: UAV Object Tracking Algorithm Based on Multi-Scale Feature

摘要: UAVs have entered various fields of life, and object tracking is one of the key technologies for UAV applications. However, there are various challenges in practical applications, such as the scale change of video images, motion blur and too high shooting angle leading to the tracked objects being too small, resulting in poor tracking accuracy. To cope with the problem that small targets are poorly tracked by UAVs due to less effective information output from the deep residual network, a SiamMFF tracking method that introduces an efficient multi-scale feature fusion strategy is proposed. The method aggregates features at different scales, and at the same time, replaces the ordinary convolution with deformable convolution to increase the sense field of convolution operation to enhance the feature extraction capability. The experimental results show that the proposed algorithm improves the success rate and accuracy of small target tracking.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001173008200001  
Indexed Date:  
2024-03-08  
Institution Address:  
Hebei Univ Sci & Technol, Sch Informat Sci & Engn, Shijiazhuang 050018, Peoples R China  
Institution Address:  
China Mobile Hebei Co Ltd, Zhangjiakou Branch, Beijing 100032, Peoples R China  
Institution Address:  
Hebei Technol Innovat Ctr Intelligent IoT, Shijiazhuang 050018, Peoples R China  
Institution:  
Institution:  
China Mobile Hebei Co Ltd  
Institution:  
Hebei Technol Innovat Ctr Intelligent IoT

### 276. Water Segmentation for Unmanned Ship Navigation Based on Multi-Scale

摘要: The segmentation of the navigation area from water images is of great significance in the safe and automated navigation of unmanned vessels. However, accurate segmentation of water boundaries in real time under the interference of water-surface light changes is still a challenging problem. In this study, a river water segmentation method based on a multi-scale feature fusion network (Fuse-ResNet), which improves the segmentation of river water boundaries by focusing more on the target features, is proposed. An Atrous Spatial Pyramid Pooling (ASPP) mechanism is introduced to expand the sensory field and enhance the relevance of contextual information at different scales. Experimental results based on the USVInland unmanned vessel dataset show that the proposed method achieves a mean Pixel Accuracy (mPA) of 99.25% and a mean Intersection over Union (mIoU) of 98.45% in the water category, which are better than those achieved by mainstream models. The interference of background noise resulting from water surface reflection can be suppressed to improve the accuracy of water segmentation for unmanned vessel navigation.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001442321100001  
Indexed Date:  
2025-03-18  
Institution Address:  
Wuhan Univ Technol, Sch Naval Architecture Ocean & Energy Power Engn, Wuhan 430070, Peoples R China  
Institution Address:  
Beibu Gulf Univ, Key Lab Beibu Gulf Offshore Engn Equipment & Techn, Qinzhou 535011, Peoples R China  
Institution:  
Institution:

### 277. Feature fusion of multi-granularity and multi-scale for facial

摘要: Although great progress has been made in facial expression recognition, it still faces challenges such as occlusion and pose changes in real-world scenario. To address this issue, we propose a simple yet effective multi-granularity and multi-scale feature fusion network (MM-Net) to achieve robust expression recognition without either manually extracting local patches or designing complex sub-networks. Specifically, we use a puzzle generator to divide the image into local regions of different granularity, which are then randomly shuffled and reorganized to form a new input image. By feeding the facial puzzles in order from fine-grained to coarse-grained, the network progressively mines the local fine-grained information, the coarse-grained information, and the global information. Besides, considering the subtle inter-class variation characteristic of different expressions, we use the multi-scale feature fusion strategy in the shallow feature extraction module to obtain global features with detailed information for capturing the subtle differences in facial expression images. Extensive experimental results on three in-the-wild FER benchmarks demonstrate the superiority of the proposed MM-Net compared to state-of-the-art methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001003038200002  
Indexed Date:  
2023-06-18  
Institution Address:  
Guangxi Normal Univ, Sch Elect & Informat Engn, Guilin 541004, Peoples R China  
Institution:

### 278. SGMFNet: a remote sensing image object detection network based on

摘要: When natural image detection methods are applied to remote sensing images, their detection performance is often unsatisfactory due to the random distribution of objects, complex backgrounds, and significant scale changes. In order to better detect objects with complex backgrounds and significant scale changes in remote sensing images, this study presents SGMFNet, a remote sensing image object detection network based on spatial global attention (SGA) and multi-scale feature fusion (MFF). The SGA inserted into the backbone network can better model context information, suppress irrelevant background, and build powerful feature information, making it easier for subsequent MFF to extract scale-invariant information from adjacent feature layers. This study evaluates the performance of SGMFNet on remote sensing datasets DIOR, NWPU VHR-10, and RSD-GOD. Quantitative and qualitative results on three datasets demonstrate the superiority of SGMFNet in remote sensing object detection and its outperformance compared with other state-of-the-art methods. Therefore, SGMFNet can assist in high-precision urban planning, military monitoring, and other tasks.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2150-704X  
eISSN:  
2150-7058  
入藏号:  
WOS:001196410400001  
Indexed Date:  
2024-04-12  
Institution Address:  
Tianjin Univ, Sch Microelect, Tianjin, Peoples R China  
Institution:

### 279. Defect Detection of Photovoltaic Modules Based on Multi-Scale Feature

摘要: A photovoltaic modules defect detection algorithm based on multi-scale feature fusion is proposed to address the challenges of complex defect backgrounds, large differences in defect scales, and a high number of small target defects that traditional object detection algorithms cannot solve. The algorithm is based on the YOLOv5s framework. Firstly, a coordinate attention mechanism is embedded in the backbone network to extract important defect shapes and enhance the network's feature extraction ability. Secondly, a bidirectional feature pyramid network is used in the Neck network to adaptively fuse image features of different scales using adaptive weights. Finally, a tiny target detection layer is added to the prediction layer, and the ASFF detection head is used to adaptively fuse different output layers to reduce the loss of target feature information. The improved algorithm is validated on a photovoltaic component dataset, and the experimental results show that it can quickly and accurately identify defects, with an mAP of 91.9% and a recall rate of 90.8%, which represents a 3.2 and 4.5 percentage points improvement in mAP and recall rate, respectively, compared to the YOLOv5s network.  
摘要:  
针对光伏组件缺陷背景复杂,缺陷尺度差异较大,小目标缺陷较多等传统目标检测算法无法解决的问题,提出了一种基于多尺度特征融合的光伏组件缺陷检测算法。算法以YOLOv5s为框架,在主干网络中嵌入坐标注意力机制,用于提取重要的缺陷形态,增强网络特征提取能力;在颈部网络中使用双向特征金字塔,以自适应权重的方式融合不同尺度的图像特征;在预测层添加微小目标检测层,并结合ASFF检测头自适应融合不同输出层,减少目标特征信息丢失。在光伏组件数据集上进行验证,实验表明改进后的算法可以快速精准识别缺陷,其中mAP达到了91.9%,召回率达到了90.8%,相比于YOLOv5s网络,mAP和召回率分别提升了3.2、4.5个百分点。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1002-8331  
入藏号:  
CSCD:7664176  
Indexed Date:  
2024-05-17  
Institution Address:  
陕西科技大学,电气与控制工程学院, 西安, 陕西 710000, 中国  
Institution Address:  
School of Electrical and Control Engineering, Shaanxi University of Science and Technology, Xi'an, Shaanxi 710000, China  
Institution:  
陕西科技大学,电气与控制工程学院  
Institution:  
School of Electrical and Control Engineering, Shaanxi University of Science and Technology

### 280. Three-Dimensional Point Cloud Object Detection Based on Feature Fusion

摘要: With the continuous emergence and development of 3D sensors in recent years, it has become increasingly convenient to collect point cloud data for 3D object detection tasks, such as the field of autonomous driving. But when using these existing methods, there are two problems that cannot be ignored: (1) The bird's eye view (BEV) is a widely used method in 3D objective detection; however, the BEV usually compresses dimensions by combined height, dimension, and channels, which makes the process of feature extraction in feature fusion more difficult. (2) Light detection and ranging (LiDAR) has a much larger effective scanning depth, which causes the sector to become sparse in deep space and the uneven distribution of point cloud data. This results in few features in the distribution of neighboring points around the key points of interest. The following is the solution proposed in this paper: (1) This paper proposes multi-scale feature fusion composed of feature maps at different levels made of Deep Layer Aggregation (DLA) and a feature fusion module for the BEV. (2) A point completion network is used to improve the prediction results by completing the feature points inside the candidate boxes in the second stage, thereby strengthening their position features. Supervised contrastive learning is applied to enhance the segmentation results, improving the discrimination capability between the foreground and background. Experiments show these new additions can achieve improvements of 2.7%, 2.4%, and 2.5%, respectively, on KITTI easy, moderate, and hard tasks. Further ablation experiments show that each addition has promising improvement over the baseline.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001192467800001  
Indexed Date:  
2024-04-03  
Institution Address:  
Xidian Univ, Collaborat Innovat Ctr Quantum Informat Shaanxi Pr, Int Res Ctr Intelligent Percept & Computat,Minist, Sch Artificial Intelligence,Joint Int Res Lab Inte, Xian 710071, Peoples R China  
Institution Address:  
54th Res Inst China Elect Technol Grp Corp, Shijiazhuang 050081, Peoples R China  
Institution:  
Institution:

### 281. Real-time detection algorithm for digital meters based on multi-scale

摘要: Aiming at the problems of insufficient feature fusion, large number of network parameters and low target saliency in the current digital meter detection algorithm, a digital meter detection algorithm based on YOLOv5s is designed. First, a new feature fusion structure Bi-Directional Feature Pyramid Network Based on Multi-Scale Feature Fusion is designed to realize the full fusion between different scale feature maps and improve the detection accuracy; second, a new convolutional module Ghostconv Combined Channel Shuffle, is designed to realize the lightweight design of the network and meet the requirements of real-time substation detection tasks; finally, to improve the network's ability to characterize the instrumented digits, the Convolutional Block Attention Module is introduced in the backbone network to further enhance the network performance. Experiments are carried out on the homemade dataset, and the experimental results show that the algorithm proposed in this paper improves the average accuracy by 2.84-98.58% compared with the original network; the amount of network parameters is reduced by 26.4%, and the detection speed is improved by 25 FPS, and the detection time for each image is only 0.012 s. Compared with other digital meter detection algorithms, the network performance and the number of parameters also have a great advantage.  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001164548400002  
Indexed Date:  
2024-03-04  
Institution Address:  
Xian Univ Sci & Technol, Xian 710054, Peoples R China  
Institution:

### 282. A Novel Multi-Scale Feature Enhancement U-Shaped Network for Pixel-Level

摘要: Timely and accurate detection of pavement cracks, the most common type of road damage, is essential for ensuring road safety. Automatic image segmentation of cracks can accurately locate their pixel positions. This paper proposes a Multi-Scale Feature Enhanced U-shaped Network (MFE-UNet) for pavement crack detection. This network model uses a Residual Detail-Enhanced Block (RDEB) instead of a conventional convolution in the encoder-decoder process. The block combines Efficient Multi-Scale Attention to enhance its feature extraction performance. The Multi-Scale Gating Feature Fusion (MGFF) is incorporated into the skip connections, enhancing the fusion of multi-scale features to capture finer crack details while maintaining rich semantic information. Furthermore, we created a pavement crack image dataset named China\_MCrack, consisting of 1500 images collected from road surfaces using smartphone-mounted motorbikes. The proposed network was trained and tested on the China\_MCrack, DeepCrack, and Crack-Forest datasets, with additional generalization experiments on the BochumCrackDataset. The results were compared with those of the U-Net model, ResUNet, and Attention U-Net. The experimental results show that the proposed MFE-UNet model achieves accuracies of 82.95%, 91.71%, and 69.02% on three datasets, namely, China\_MCrack, DeepCrack, and Crack-Forest datasets, respectively, and the F1\_score is improved by 1-4% compared with other networks. Experimental results demonstrate that the proposed method is effective in detecting cracks at the pixel level.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001364361900001  
Indexed Date:  
2024-12-03  
Institution Address:  
Shandong Jiaotong Univ, Sch Rail Transportat, Jinan 250357, Peoples R China  
Institution Address:  
Shandong Prov Transportat Ind, Key Lab Rail Transit Safety Technol & Equipment, Jinan 250357, Peoples R China  
Institution:  
Institution:  
Shandong Prov Transportat Ind

### 283. DMFF-YOLO: YOLOv8 Based on Dynamic Multiscale Feature Fusion for Object

摘要: With the rapid proliferation of drones across various domains, aerial target detection has become increasingly crucial. However, the targets in aerial images present challenges such as scale variation, small size, and density, leading to suboptimal performance of current detectors on aerial images. Based on the aforementioned challenges, we design an efficient aerial target detection algorithm called DMFF-YOLO. Specifically, to address the issues of small target size and scale variation, we design the DMFF neck structure, adding a small target detection head to tackle the small target size problem, using the DMC module to fuse different scale features for enriching detailed information, and employing the DSSFF module to construct a scale sequence space to solve the target scale variation problem. In the network backbone, we employ RFCBAMConv modules as downsampling layers, which interact with receptive-field features to mitigate the information disparity caused by positional changes and outperform traditional convolutional layers. Finally, we design the Soft-NMS-CIoU module to address the issue of suppressing adjacent boxes due to dense targets. On the VisDrone dataset, compared to the original algorithm, our method reduces the number of parameters by 31.1% while achieving an 11.7% improvement in mAP50. Extensive experiments on the VisDrone, DOTA, and UAVDT datasets demonstrate that the proposed algorithm performs well in aerial image detection tasks.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001316072800001  
Indexed Date:  
2024-09-25  
Institution Address:  
China West Normal Univ, Sch Elect Informat Engn, Nanchong 637009, Peoples R China  
Institution:

### 284. AlgaeClass\_Net: Optimizing Few-Shot Marine Microalgae Classification

摘要: As the eutrophication of the water body becomes more and more serious, the algae in the water body grow in large quantities and eventually form harmful algal blooms, causing great harm to the marine ecosystem. Therefore, how to quickly and accurately identify algae and make precautions becomes the key to solving this problem. Currently, more than tens of thousands of microalgae species are known around the world, but publicly available data are sparse, many of the species are characterized similarly to each other, and it is currently challenging to train an effective classification model with limited data. Existing few-shot learning classification algorithms that utilize meta-learning for metrics can be a good solution to this problem. In this paper, an AlgaeClass\_Net algorithm that combines an improved multi-scale feature fusion with a feature enhancement module is proposed for the fine-grained features of microalgae. Furthermore, it utilizes a metric learning approach to classify few-shot microalgae by calculating the distances between the feature vectors of samples in the query set and the feature vectors of samples in the support set. The experimental results showed that the method achieves 78.55% and 91.20% classification accuracies under different tasks of 5-way 1-shot and 5-way 5-shot, respectively, with an improvement of 3.51% and 4.97% on the suboptimal model, respectively. It provides new research ideas for the identification of marine microalgae and the development and utilization of marine renewable energy.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001410383800017  
Indexed Date:  
2025-02-09  
Institution Address:  
Dalian Ocean Univ, Coll Informat Engn, Dalian 116000, Peoples R China  
Institution Address:  
Hangzhou Yunxi Smart Vis Technol Co Ltd, Dalian 116000, Peoples R China  
Institution Address:  
Ningbo Inst Dalian Univ Technol, Ningbo 315000, Peoples R China  
Institution Address:  
Dalian Ocean Univ, Key Lab Environm Controlled Aquaculture, Minist Educ, Dalian 116000, Peoples R China  
Institution:  
Institution:  
Hangzhou Yunxi Smart Vis Technol Co Ltd  
Institution:  
Ningbo Inst Dalian Univ Technol  
Institution:

### 285. Improved DeepLabv3+ Model for Surface Defect Detection on Steel Plates

摘要: To address problems of rough edge segmentation, missed detection and high false detection rate in steel surface defect detection, a multi- scale feature fusion detection method based on DeepLabv3+ with attention mechanism is proposed. Firstly, in the decoding region of the DeepLabv3+ network, the multi-scale feature information is fully utilized and the leap-frog feature fusion is optimized to retain the shallow features, while a more refined up-sampling operation on the deep features is performed to obtain finer defect edges. Secondly, a coordinate attention mechanism is introduced into the coding region backbone network ResNet101 to enhance the capability of the feature extraction and improve the segmentation accuracy. In addition, an optimized loss function combining weighted Dice loss and binary cross entropy loss(BCEloss) is designed to alleviate the problem of sample imbalance and improve segmentation accuracy. The Dice coefficient and mIoU values of the advanced DeepLabv3+ network are improved by 6.0% and 7.92% respectively, with more accurate edge segmentation of scratch defects and significant improvement in the segmentation of pits, edge cracks and iron oxide defects. The experimental results validate the effectiveness of the method in dealing with the steel surface defect problem.  
摘要:  
针对钢板表面缺陷检测中存在的边缘分割粗糙、漏检和误检率高等问题,提出了一种引入注意力机制的多尺度特征融合的DeepLabv3+检测方法。在DeepLabv3+网络的解码区中,充分利用多尺度特征信息,对跃层特征融合进行优化,保留浅层特征并对深层特征进行了细化的上采样操作,获得更精细的缺陷边缘;在编码区主干网络ResNet101中引入坐标注意力机制,增强特征提取能力,提高分割准确率。设计了加权Dice损失和二元交叉熵损失(BCEloss)结合的优化损失函数来缓解样本不均衡的问题,提高分割精度。改进DeepLabv3+网络的Dice系数和mIoU值分别提高了6.0%和7.92%,刮痕缺陷边缘分割更准确,对凹坑、边缘裂纹与氧化铁皮缺陷的分割效果提升明显,实验结果验证了该方法处理钢板表面缺陷问题的有效性。  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1002-8331  
入藏号:  
CSCD:7549184  
Indexed Date:  
2023-12-16  
Institution Address:  
北京信息科技大学计算机学院, 北京 100101, 中国  
Institution Address:  
北京信息科技大学信息管理学院, 北京 100192, 中国  
Institution Address:  
Computer School, Beijing Information Science and Technology University, Beijing 100101, China  
Institution Address:  
School of Information Management, Beijing Information Science and Technology University, Beijing 100192, China  
Institution:  
北京信息科技大学计算机学院  
Institution:  
北京信息科技大学信息管理学院  
Institution:  
Computer School, Beijing Information Science and Technology University  
Institution:  
School of Information Management, Beijing Information Science and Technology University

### 286. Iterative Fusion and Dual Enhancement for Accurate and Efficient Object

摘要: Single Shot Multibox Detector (SSD) uses multi-scale feature maps to detect and recognize objects, which considers the advantages of both accuracy and speed, but it is still limited to detecting small-sized objects. Many researchers design new detectors to improve the accuracy by changing the structure of the multi-scale feature pyramid which has proved very useful. But most of them only simply merge several feature maps without making full use of the close connection between features with different scales. In contrast, a novel feature fusion module and an effective feature enhancement module is proposed, which can significantly improve the performance of the original SSD. In the feature fusion module, the feature pyramid is produced through iteratively fusing three feature maps with different receptive fields to obtain contextual information. In the feature enhancement module, the features are enhanced along the channel and spatial dimensions at the same time to improve their expression ability. Our network can achieve 82.5% mean Average Precision (mAP) on the VOC 2007 test, 81.4% mAP on the VOC 2012 test and 34.8% mAP on COCO test-dev2017, respectively, with the input size 512 x 512. Comparative experiments prove that our method outperforms many state-of-the-art detectors in both aspects of accuracy and speed.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0218-1266  
eISSN:  
1793-6454  
入藏号:  
WOS:000957007800002  
Indexed Date:  
2023-04-22  
Institution Address:  
Hefei Univ Technol, Sch Comp Sci & Informat, Key Lab Knowledge Engn Big Data, Minist Educ, Hefei 230000, Peoples R China  
Institution Address:  
Anhui Med Univ, Hosp 2, Hefei 230000, Peoples R China  
Institution:  
Institution:

### 287. AFFD-Net: A Dual-Decoder Network Based on Attention-Enhancing and

摘要: The morphological characteristics of retinal vessels in the fundus serve as the primary basis for diagnosing and assessing the risk of ophthalmic diseases. An effective segmentation scheme for retinal vessels can aid in the early diagnosis and treatment of these diseases, as well as help prevent their progression. However, accurate vessel segmentation is challenging due to the low contrast of fundus images and the complexity of the vessels' morphological structure. To address the low sensitivity and poor generalization ability of the existing methods in vascular extraction, a Dual-decoder Network based on Attention-enhancing and multi-scale Feature Fusion (AFFD-Net) is proposed. AFFD-Net inherits the codec concept of U-Net. To improve the performance of our U-Net model, we made two modifications. Firstly, we reduced the number of convolution kernel filters in each layer, thereby significantly reducing the number of training parameters. This helps to avoid overfitting and improves the model's ability to generalize. Secondly, we added a Multi-scale Feature Extraction (MFE) module and an M/A intermediate decoder to enhance the model's sensitivity. MFE is designed as the first encoding unit of AFFD-Net to obtain rich vascular features in the complex anatomical background and adapt to the large-scale variations of vessels. The M/A intermediate decoder is composed of the Multi-scale Feature Fusion (MFF) module and the Attention-enhancing Hybrid Feature Fusion (AHFF) module. The MFF module integrates deep semantic information and shallow spatial information to ensure that the features at each scale in the middle layer are fully utilized. The AHFF module adaptively fuses the hybrid features at different scales to generate two feature descriptors with different focuses which can improve the expressiveness of the model. AFFD-Net is evaluated on three public databases including DRIVE, STARE, and CHASE\_DB1, and the sensitivity values obtained are 84.19%, 84.58%, and 82.62%, respectively. It has higher sensitivity and better generalization ability than other state-of-the-art methods. Compared with classical networks including U-Net, U-Net++, and U-Net3+, AFFD-Net has fewer parameters and higher segmentation accuracy. Our proposed segmentation model exhibits superior performance across a range of metrics, indicating its promising potential for practical applications.  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
4  
ISSN:  
2169-3536  
入藏号:  
WOS:000991571500001  
Indexed Date:  
2023-06-07  
Institution Address:  
Changchun Univ Sci & Technol, Sch Life Sci & Technol, Changchun 130022, Peoples R China  
Institution Address:  
RMIT Univ, Dept Informat Syst & Business Analyt, Melbourne, Vic 3001, Australia  
Institution Address:  
Univ Melbourne, Sch Comp & Informat Syst, Parkville, Vic 3010, Australia  
Institution Address:  
Monash Univ, Dept Software Syst & Cybersecur, Clayton, Vic 3800, Australia  
Institution Address:  
Jilin Agr Sci & Technol Univ, Sch Elect & Informat Engn, Jilin 132109, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 288. Synthetic Aperture Radar Ship Detection Based on Efficient Multiscale

摘要: Due to the scattering effect of synthetic aperture radar (SAR) and the resolution limitation of wavelength and antenna size, it is difficult to obtain the details and boundary information of small-sized targets, which affects the detection accuracy. To improve the accuracy of SAR ship detection and reduce the false detection rate, SAR ship detection based on an efficient multiscale feature enhancement network (EMFENet) is proposed. First, in the backbone network, a multifeature fusion structure is used to construct a step-by-step residual partial convolution block as the main feature extraction module, which improves the recognition ability of ship targets and effectively reduces the number of parameters. Second, the feature fusion part constructs an efficient layer convolution block with a spatial channel attention mechanism and an aggregation structure and pays more attention to the geometric information of the model. The multiscale fusion feature pyramid network is constructed using the receptive field block, and the multiscale features are further fused to enhance the feature capture ability of the overall structure of the network. Finally, the normalized Gaussian Wasserstein distance metric is introduced into the detection head loss regression function to enhance the detection ability of small targets. The effectiveness of the proposed network was compared with mainstream detection algorithms using the high-resolution SAR images dataset (HRSID) and the SAR ship detection dataset (SSDD). Experiments showed that the proposed network has significant advantages in improving the detection accuracy of SAR ships and in reducing false and missed detections.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0018-9251  
eISSN:  
1557-9603  
入藏号:  
WOS:001373839100002  
Indexed Date:  
2024-12-19  
Institution Address:  
Wuxi Univ, Wuxi 214105, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Elect & Informat Engn, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Nanjing 210044, Peoples R China  
Institution:  
Institution:  
Institution:

### 289. MFEFNet: A Multi-Scale Feature Information Extraction and Fusion Network

摘要: Unmanned aerial vehicles (UAVs) are now widely used in many fields. Due to the randomness of UAV flight height and shooting angle, UAV images usually have the following characteristics: many small objects, large changes in object scale, and complex background. Therefore, object detection in UAV aerial images is a very challenging task. To address the challenges posed by these characteristics, this paper proposes a novel UAV image object detection method based on global feature aggregation and context feature extraction named the multi-scale feature information extraction and fusion network (MFEFNet). Specifically, first of all, to extract the feature information of objects more effectively from complex backgrounds, we propose an efficient spatial information extraction (SIEM) module, which combines residual connection to build long-distance feature dependencies and effectively extracts the most useful feature information by building contextual feature relations around objects. Secondly, to improve the feature fusion efficiency and reduce the burden brought by redundant feature fusion networks, we propose a global aggregation progressive feature fusion network (GAFN). This network adopts a three-level adaptive feature fusion method, which can adaptively fuse multi-scale features according to the importance of different feature layers and reduce unnecessary intermediate redundant features by utilizing the adaptive feature fusion module (AFFM). Furthermore, we use the MPDIoU loss function as the bounding-box regression loss function, which not only enhances model robustness to noise but also simplifies the calculation process and improves the final detection efficiency. Finally, the proposed MFEFNet was tested on VisDrone and UAVDT datasets, and the mAP0.5 value increased by 2.7% and 2.2%, respectively.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
被引频次合计:  
8  
eISSN:  
2504-446X  
入藏号:  
WOS:001232295900001  
Indexed Date:  
2024-06-02  
Institution Address:  
Henan Univ, Henan Key Lab Big Data Anal & Proc, Kaifeng 475000, Peoples R China  
Institution Address:  
Henan Univ, Sch Comp & Informat Engn, Kaifeng 475000, Peoples R China  
Institution:  
Institution:

### 290. Crowd Counting based on Multi-level Multi-scale Feature

摘要: Crowd counting has drawn more and more attention for its significance in reality application. However, it's still a challenging task because of scale variation in images. In this paper, we propose a model to extract and refine features with abundant scale-relevant information, which consists of Multi-layer Multi-scale Feature Extraction Network (MLMS) and Dependency-based Feature Fusion Network (DFF). MLMS plays a role as feature extractor. Three multi-scale feature extraction modules (MSFE) are designed with dilated convolution layers and inserted in different levels of MLMS, which improve the ability for multi-scale feature extraction. DFF plays a role as feature refiner. DFF explores the dependency between hierarchical features. It's the first time in crowd counting to use Long-short term memory (LSTM) to filter information and fuse the features with the assistance of the dependency. Our model provides new ideas for solving scale-relevant problems from two angels: scale feature extraction and fusion. In this way, our model extracts scale-relevant features and refines the features further. Experiments on four challenging datasets ShanghaiTech Part A/B, UCF\_QNRF and UCF\_CC\_50, getting Mean Absolute Error (MAE) 65.3/8.3/113.2/216.3, demonstrate the effectiveness of the proposed model.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:001007615100001  
Indexed Date:  
2023-06-27  
Institution Address:  
Beijing Inst Technol, Sch Informat & Elect, Beijing, Peoples R China  
Institution:

### 291. MFDNN: multi-scale feature-weighted dual-neck network for underwater

摘要: With the rapid development of object detection technology, underwater object detection has gradually become a hot topic. Due to the complex underwater environment, some object detection algorithms still encounter difficulties in detecting targets similar to the background. For the above problems, we propose a multi-scale feature-weighted dual-neck network (MFDNN) for underwater object detection. Our contribution is mainly divided into three parts. First, an enhanced feature extraction network, namely the dual-neck network, is designed to process and reuse the features extracted from the backbone network. Second, an attention mechanism is embedded in one of the neck networks to reweight features and pay more attention to important features. In addition, we introduce the adaptively spatial feature fusion mechanism to adaptively weight the features extracted at multiple scales. As demonstrated in comprehensive experiments, the mean average precision of our MFDNN can reach 87.79% and 86.51% on the underwater datasets URPC2019 and URPC2020, respectively.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:001286842600005  
Indexed Date:  
2024-08-16  
Institution Address:  
Dalian Maritime Univ, Sch Marine Elect Engn, Dalian, Peoples R China  
Institution:

### 292. Object Detection Algorithm in Remote Sensing Images Based on Improved

摘要: Remote sensing target detection is an important aspect in the fields of environmental monitoring and circuit patrol. A remote sensing target detection algorithm based on YOLOX is proposed for the difficulties of remote sensing images with large target scale differences, blurred targets and high background complexity. First, a regional context aggregation module is proposed to expand the perceptual field using the dalited convolutions with different expansion rates to obtain multi-scale contextual information, which is beneficial to the detection of the small targets. Second, the feature fusion module is proposed, and two different scale transformation modules are used to achieve the fusion of features at different scales, fully fusing shallow location information with deep semantic information to improve the detection performance of the network for targets at different scales. Finally, a feature enhancement module is introduced to the multiscale feature fusion network part and combined with the attention mechanism CAS [CA (coordinate attention) with SimAM (simple parameter-free attention module)] to make the network pay more attention to the target information and ignore the interference of complex background, while the shallow feature layer is fused with the deep detection layer for feature fusion to prevent the low detection performance affected by the loss of feature information at the prediction end. The experimental results show that the improved algorithm achieves 73. 87% and 96. 22% detection accuracy on DIOR and RSOD remote sensing datasets, which is 4. 08 and 1. 34 percentage points higher than the original YOLOX algorithm, and has superiority in both detection accuracy and detection speed compared with other advanced algorithms.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1006-4125  
入藏号:  
WOS:001284198700030  
Indexed Date:  
2024-09-18  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Elect & Informat Engn, Nanjing 210044, Jiangsu, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Jiangsu Collaborat Innovat Ctr Atmospher Environm, Nanjing 210044, Jiangsu, Peoples R China  
Institution:  
Institution:

### 293. Fittings Detection Method Based on Multi-Scale Geometric Transformation

摘要: Overhead transmission lines are important lifelines in power systems, and the research and application of their intelligent patrol technology is one of the key technologies for building smart grids. The main reason for the low detection performance of fittings is the wide range of some fittings' scale and large geometric changes. In this paper, we propose a fittings detection method based on multi-scale geometric transformation and attention-masking mechanism. Firstly, we design a multi-view geometric transformation enhancement strategy, which models geometric transformation as a combination of multiple homomorphic images to obtain image features from multiple views. Then, we introduce an efficient multiscale feature fusion method to improve the detection performance of the model for targets with different scales. Finally, we introduce an attention-masking mechanism to reduce the computational burden of model-learning multiscale features, thereby further improving model performance. In this paper, experiments have been conducted on different datasets, and the experimental results show that the proposed method greatly improves the detection accuracy of transmission line fittings.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:000998002900001  
PubMed ID:  
37430837  
Indexed Date:  
2023-06-09  
Institution Address:  
CSG EHV Power Transmiss Co, Operat & Maintenance Ctr Informat & Commun, Guangzhou 510000, Peoples R China  
Institution Address:  
North China Elect Power Univ, Dept Elect & Commun Engn, Baoding 071003, Peoples R China  
Institution:  
CSG EHV Power Transmiss Co  
Institution:

### 294. Feedback Network for Compact Thin Cloud Removal

摘要: The thin cloud removal (CR) technique has great practical value for the application of remote-sensing images. Existing deep-learning-based methods have attained remarkable achievements. However, most of them neglect the inherent feature correlations in deeper layers due to learning successively. In this letter, we propose a compact thin CR network based on the feedback (FB) mechanism, called CRFB-Net, which leverages the high-level features as FB information to modulate shallow representations. CRFB-Net employs the recurrent architecture to achieve such an FB scheme. Specifically, the restoration process does not terminate after obtaining an output. In this case, the output of intermediate iterations will flow into the next iteration as FB. For better utilization of FB, a multiscale feature fusion block (MFFB) is designed to refine the low-level representations from three scales. Furthermore, we introduce a curriculum learning (CL) strategy to train the CRFB-Net by gradually increasing the complexity of restoration, through which a sharper result is produced step by step. Extensive experiments demonstrate the superiority of our CRFB-Net, outperforming state-of-the-art (SOTA).  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000967174700001  
Indexed Date:  
2023-04-24  
Institution Address:  
Beihang Univ, Sch Astronaut, Dept Aerosp Informat Engn, Beijing 100191, Peoples R China  
Institution Address:  
Shanghai Aerosp Control Technol Inst, Shanghai 201109, Peoples R China  
Institution:  
Institution:  
Shanghai Aerosp Control Technol Inst

### 295. Multi-CrackNet: A Fast Segmentation and Quantification Combined Method

摘要: Discontinuity investigation and characterization onsite is a labor-dependent work because current techniques cannot precisely handle multiple discontinuity identifications automatically under different work conditions. This paper proposes the multi-CrackNet which enables us to identify and segment linear discontinuities (joints and cracks) for random types of rock surface. A modified feature extraction network called the multiscale feature fusion pyramid network (MFFPN) has been developed based on FPN to capture and fuse more sensitive texture features of cracks across different types of background. With the help of a new training scheme by setting up 3 stages of training to simulate the human-based learning process, the established model can learn more features steadily and robustly from well-labelled databases. Additionally, a hybrid pixel-level quantification method is proposed to automatically compute the length, width, and inclination of cracks. Results show that the proposed method can achieve a detection accuracy of 87.1% for 1 to 9 sets of cracks on the rock surface across different types of rock. Case studies in Anshan West are provided to verify the reliability and accuracy of our method in macrolinear discontinuity identification and quantification, which sees great potentials in site investigation by saving a large amount of labor force.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
4  
ISSN:  
1468-8115  
eISSN:  
1468-8123  
入藏号:  
WOS:000888552300001  
Indexed Date:  
2022-12-03  
Institution Address:  
Univ Sci & Technol Beijing, Dept Civil Engn, Beijing 100083, Peoples R China  
Institution Address:  
Univ Sci & Technol Beijing, Beijing Key Lab Urban Underground Space Engn, Beijing 100083, Peoples R China  
Institution:  
Institution:

### 296. Hierarchical loop closure detection with weighted local patch features

摘要: Maintaining high-precision localization and ensuring map consistency are crucial objectives for mobile robots. However, loop closure detection remains a challenging aspect of their operation because of viewpoint and appearance changes. To address this issue, this paper proposes WP-VLAD, a novel hierarchical loop closure detection method that tightly couples global features and weighted local patch-level features (WPs). WP-VLAD employs MobileNetV3 as the backbone network for feature extraction, and integrates a trainable vector of local aggregated descriptors (VLAD) for compact global and local feature representation. A hierarchical navigable small world method is used to retrieve loop candidate frames based on the global features, whereas a multiscale feature fusion weighted map prediction module assigns weights to the local patches during mutual nearest neighbour matching. The proposed weight allocation strategy emphasizes salient regions, reducing interference from dynamic objects. The experimental results on benchmark datasets demonstrate that WP-VLAD significantly improves matching performance while maintaining efficient computation, exhibiting strong generalizability and robustness across various complex environments.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:001389428600005  
Indexed Date:  
2025-01-09  
Institution Address:  
Beijing Univ Technol, Sch Informat Sci & Technol, Beijing 100124, Peoples R China  
Institution:

### 297. YOLO-FGD: a fast lightweight PCB defect method based on FasterNet and

摘要: With the rapid expansion of the electronics industry, the demand for high-quality printed circuit boards has surged. However, existing PCB defect detection methods suffer from various limitations, such as slow speeds, low accuracy, and restricted detection scope, often leading to false positives and negatives. To overcome these challenges, this paper presents YOLO-FGD, a novel detection model. YOLO-FGD replaces YOLOv5's backbone network with FasterNet, significantly accelerating feature extraction. The Neck section adopts the Gather-and-Distribute mechanism, which enhances multiscale feature fusion for small targets through convolution and self-attention mechanisms. Integration of the C3\_Faster feature extraction module effectively reduces the number of parameters and the number of FLOPs, accelerating the computations. Experiments on the PCB-DATASETS dataset show promising results: the mean average precision50 reaches 98.8%, the mean average precision50-95 reaches 57.2%, the computational load is reduced to 11.5 GFLOPs, and the model size is only 12.6 MB, meeting lightweight standards. These findings underscore the effectiveness of YOLO-FGD in efficiently detecting PCB defects, providing robust support for electronic manufacturing quality control.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001261579900001  
Indexed Date:  
2024-07-12  
Institution Address:  
Chengdu Univ Technol, Coll Mech & Elect Engn, Chengdu 610059, Sichuan, Peoples R China  
Institution:

### 298. CRS-Diff: Controllable Remote Sensing Image Generation With Diffusion

摘要: The emergence of generative models has revolutionized the field of remote sensing (RS) image generation. Despite generating high-quality images, existing methods are limited in relying mainly on text control conditions, and thus do not always generate images accurately and stably. In this article, we propose CRS-Diff, a new RS generative framework specifically tailored for RS image generation, leveraging the inherent advantages of diffusion models while integrating more advanced control mechanisms. Specifically, CRS-Diff can simultaneously support text-condition, metadata-condition, and image-condition control inputs, thus enabling more precise control to refine the generation process. To effectively integrate multiple condition control information, we introduce a new conditional control mechanism to achieve multiscale feature fusion (FF), thus enhancing the guiding effect of control conditions. To the best of our knowledge, CRS-Diff is the first multiple-condition controllable RS generative model. Experimental results in single-condition and multiple-condition cases have demonstrated the superior ability of our CRS-Diff to generate RS images both quantitatively and qualitatively compared with previous methods. Additionally, our CRS-Diff can serve as a data engine that generates high-quality training data for downstream tasks, e.g., road extraction. The code is available at https://github.com/Sonettoo/CRS-Diff.  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
4  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001313400800027  
Indexed Date:  
2024-09-25  
Institution Address:  
Xi An Jiao Tong Univ, Sch Comp Sci & Technol, Xian 710049, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Key Lab Intelligent Networks & Network Secur, Minist Educ, Xian 710049, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Sch Informat & Commun Engn, Xian 710049, Shaanxi, Peoples R China  
Institution Address:  
Xidian Univ, Sch Cyber Engn, Xian 710049, Shaanxi, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Sch Math & Stat, Xian 710049, Shaanxi, Peoples R China  
Institution Address:  
Macau Univ Scienceand Technol, Macao Inst Syst Engn, Taipa, Macao, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Macau Univ Scienceand Technol

### 299. ECFNet: A Siamese Network With Fewer FPs and Fewer FNs for Change

摘要: High-resolution remote-sensing image change detection plays an important role in areas such as land and resources investigation, natural disaster prediction, and military strategy research. Current change detection methods often focus on extracting more discriminative features, while ignoring the information loss and imbalance problems in the process of feature fusion, which results in weakness in small change objects and edge pixels of change objects. In this letter, a simple but efficient network architecture, extraction, comparison and fusion network (ECFNet), for change detection in remote-sensing images is proposed. By constraining the number of feature channels in the fusion process by the feature comparison module (FCM), ECFNet can better utilize the fine-grained information in the multiscale feature map for result prediction, which not only improves the detection performance of small objects, but also reduces the false detection around the edge pixels of change objects. Experiments on the deeply supervised image fusion network for change detection (DSIFN-CD) test set show that ECFNet achieves state-of-the-art results with a small amount of computation.  
Web of Science 核心合集中的 "被引频次":  
28  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
29  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000966553600001  
Indexed Date:  
2023-05-22  
Institution Address:  
Xi An Jiao Tong Univ, Sch Software Engn, Xian 710049, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Coll Artificial Intelligence, Xian 710049, Peoples R China  
Institution:  
Institution:

### 300. Segmentation algorithm of lightweight bladder cancer MRI images based on

摘要: There exists serveral challenges in MRI images of bladder cancer,such as unclear tumor boundaries,small tumor areas,and discontinuous tumor distribution.The existing segmentation algorithms have huge parameters and complex calculations,and the accuracy and efficiency of the existing methods for bladder tumor segmentation need to be improved.In response to the above issues,a lightweight bladder cancer segmentation algorithm based on multiscale feature fusion(PylNet)was proposed in this paper.The algorithm can extract information of different scales through the multi-scale semantic feature extraction module.This module was designed in the coding stage to ensure the reliability and comprehensiveness of the extraction of information on small tumor regions.At the same time,the fusion module designed in another stage can quickly complete the tumor region segmentation,and the amount of parameters used is much less.Experimental results show that compared with FCN8s,SegNet,U-Net and other algorithms,the segmentation accuracy of this algorithm is improved to a certain extent where DSC reaches 88.40%,and the parameter amount is 1/13 of FCN8s,which a fast bladder MRI segmentation is achieved.  
摘要:  
膀胱癌MRI图像存在肿瘤边界不清晰、肿瘤区域较小、肿瘤分布不连续等问题,现有的分割算法参数量庞大,计算复杂,且分割精度有待提高。因此,设计了一种多尺度特征融合的轻量化膀胱癌分割算法(pyramidal convolution lightweight network,PylNet),该算法在编码阶段设计的多尺度语义特征提取模块可提取不同尺度的肿瘤区域信息,确保对微小肿瘤信息提取的可靠性和全面性;在解码阶段设计的融合模块可以在保证分割精度的同时,极大地减少算法参数量和复杂度。实验结果表明,相较于FCN8s、DeepLabV3+、U-Net等算法,PylNet算法分割精度有一定的提高,Dice系数达88.40%,参数量是FCN8s的1/13,可实现对膀胱MRI的快速分割。  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1672-4291  
入藏号:  
CSCD:7204018  
Indexed Date:  
2022-07-15  
Institution Address:  
山东中医药大学智能与信息工程学院, 济南, 山东 250355, 中国  
Institution Address:  
解放军第960医院医学工程科, 济南, 山东 250031, 中国  
Institution Address:  
山东中医药大学医学人工智能研究中心;;山东中医药大学第一临床医学院, ;;, 青岛;;济南, 山东;;山东 266112;;250355, 中国  
Institution Address:  
山东中医药大学医学人工智能研究中心;;山东中医药大学青岛中医药科学院, ;;, 青岛;;青岛, 山东;;山东 266112;;266112, 中国  
Institution Address:  
山东青年政治学院信息工程学院, 济南, 山东 250103, 中国  
Institution Address:  
College of Intelligence and Information Engineering,Shandong University of Traditional Chinese Medicine, Jinan, Shandong 250355, China  
Institution Address:  
Medical Engineering Department,the 960th Hospital of the PLA, Jinan, Shandong 250031, China  
Institution Address:  
Center for Medical Artificial Intelligence,Shandong University of Traditional Chinese Medicine;;First School of Clinical Medicine,Shandong University of Traditional Chinese Medicine, ;;, Qingdao;;Jinan, Shandong;;Shandong 266112;;250355  
Institution Address:  
Center for Medical Artificial Intelligence,Shandong University of Traditional Chinese Medicine;;Qingdao Academy of Chinese Medical Sciences,Shandong University of Traditional Chinese Medicine, ;;, Qingdao;;Qingdao, Shandong;;Shandong 266112;;266112  
Institution Address:  
School of Information Engineering,Shandong Youth University of Political Science, Jinan, Shandong 250103, China  
Institution:  
山东中医药大学智能与信息工程学院  
Institution:  
解放军第960医院医学工程科  
Institution:  
山东中医药大学医学人工智能研究中心;;山东中医药大学第一临床医学院  
Institution:  
山东中医药大学医学人工智能研究中心;;山东中医药大学青岛中医药科学院  
Institution:  
山东青年政治学院信息工程学院  
Institution:  
College of Intelligence and Information Engineering,Shandong University of Traditional Chinese Medicine  
Institution:  
Medical Engineering Department,the 960th Hospital of the PLA  
Institution:  
Center for Medical Artificial Intelligence,Shandong University of Traditional Chinese Medicine;;First School of Clinical Medicine,Shandong University of Traditional Chinese Medicine  
Institution:  
Center for Medical Artificial Intelligence,Shandong University of Traditional Chinese Medicine;;Qingdao Academy of Chinese Medical Sciences,Shandong University of Traditional Chinese Medicine  
Institution:  
School of Information Engineering,Shandong Youth University of Political Science

### 301. Multi-scale saliency features fusion model for person re-identification

摘要: Person re-identification mainly uses computer vision technology to determine whether there are specific pedestrians in the image or video. Belong to cross-device retrieval images, due to the changing style of the device led to more difficult person re-identification. The current algorithms use multi-feature fusion methods such as posture detection to match, ignoring the impact of different perspectives, postures and backgrounds on features. This paper proposes a multi-scale feature method based on saliency model. Which uses the salient image extraction algorithm to better filter out the interference of complex background parts, and uses the feature weighting method to fuse the global features and local features to achieve more robust features. Experimental results on three datasets show that the proposed method is superior to the existing method.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:000907916500001  
Indexed Date:  
2023-01-20  
Institution Address:  
Xian Univ Technol, Sch Printing Packaging & Digital Media, Xian, Shaanxi, Peoples R China  
Institution Address:  
Printing & Packaging Engn Technol Res Ctr Shaanxi, Xian, Peoples R China  
Institution:  
Institution:  
Printing & Packaging Engn Technol Res Ctr Shaanxi

### 302. Soft Pseudo-Label and Multi-Scale Feature Fusion for Person

摘要: The traditional unsupervised domain adaptive person re identification algorithm suppressed the noise of pseudo label poorly and lack inter domain generalization ability. For the above problems, an unsupervised domain adaptive person re-identification algorithm was proposed which based on soft pseudo-label and multi-scale feature reconstruction. In order to suppress pseudo-label noise, the predicted value of the parallel network is used as the soft tag, and pseudo-label noise is corrected by cross-proofreading methods, which provides a more robust soft false tag for unsupervised domain adaptive tasks. In order to enhance the generalization ability between domains, multi-scale feature reconstruction and Hadamard product feature fusion methods are used to process the deep and shallow feature layer information , realize the style conversion from source domain data to target domain and solve the problem of poor adaptability of residual network domain with instance normalization and batch normalization network, so as to enhance the generalization ability of the network to source domain and target domain. Experimental results show that the proposed algorithm has achieved good performance in both Market to Duke and Duke to Market unsupervised domain adaptive tasks, which is significantly better than the related algorithms.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1006-4125  
入藏号:  
WOS:000892704400015  
Indexed Date:  
2022-12-18  
Institution Address:  
Inner Mongolia Univ Sci & Technol, Sch Informat Engn, Baotou 014010, Inner Mongolia, Peoples R China  
Institution Address:  
Mongolia Ind Univ, Sch Informat Engn, Hohhot 010051, Inner Mongolia, Peoples R China  
Institution Address:  
Inner Mongolia Key Lab Patten Recognit & Intellig, Baotou 014010, Inner Mongolia, Peoples R China  
Institution:  
Institution:  
Mongolia Ind Univ  
Institution:  
Inner Mongolia Key Lab Patten Recognit & Intellig

### 303. High-Frequency Dual-Branch Network for Steel Small Defect Detection

摘要: Strip surface defect detection is pivotal in the steel industry for improving strip production quality. However, there is still a big gap between the existing working and the detection of small defects in strip steel in practical applications. In this paper, we propose the SSD-YOLO model, which is designed specifically for detecting small defects on strip steel surfaces. Given the challenge of feature extraction due to the small defect size, it utilizes a dual-branch feature extraction and channel-level feature fusion to enhance the expression capability of small defects. Moreover, it integrates a multiscale high-resolution detection module to achieve precise segmentation, thereby improving the overall detection accuracy of the model. The experimental results illustrate that the SSD-YOLO model, as proposed, attains a 98.0% mean average precision (mAP) and operates at 66 frames per second (FPS) when evaluated on the SSDD (Steel Small Defect Dataset). In comparison with YoloV8s, the SSD-YOLO achieves a significant improvement in accuracy, with an increase of 19.9%. The inference time and performance of our SSD-YOLO is well balanced, making it suitable for real-world deployment.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2193-567X  
eISSN:  
2191-4281  
入藏号:  
WOS:001285487600002  
Indexed Date:  
2024-08-13  
Institution Address:  
North China Univ Sci & Technol, Coll Artificial Intelligence, 21 Bohai Rd,Caofeidian Xincheng, Tangshan 063210, Hebei, Peoples R China  
Institution Address:  
North China Univ Sci & Technol, Hebei Key Lab Ind Intelligent Percept, 21 Bohai Rd,Caofeidian Xincheng, Tangshan 063210, Hebei, Peoples R China  
Institution Address:  
North China Univ Sci & Technol, Coll Elect Engn, 21 Bohai Rd,Caofeidian Xincheng, Tangshan 063210, Hebei, Peoples R China  
Institution:  
Institution:  
Institution:

### 304. Automatic recognition of pavement cracks from combined GPR B-scan and

摘要: Pavement crack detection is critical for transportation infrastructure assessment using ground penetrating radar (GPR). This paper describes a YOLOv3 model with four-scale detection layers (FDL) to detect combined B-scan and C-scan GPR images subject to poor detection effects and a high missed detection rate of small crack feature sizes. Multiscale fusion structures, efficient intersection over union (EIoU) loss function, K-means++ clustering, and hyperparameter optimization were used in this proposed model to further improve detection performance. Results indicated that the F1 score and mAP of the YOLOv3-FDL model reached 88.1% and 87.8% and had an 8.8% and 7.5% improvement on the GPR dataset of concealed cracks, respectively, compared with the YOLOv3 model. This illustrated that this model solved the problem of missed crack detection to some extent. Future studies can take these results further, especially the three-dimensional feature analysis of pavement cracks.  
Web of Science 核心合集中的 "被引频次":  
102  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
在 BIOSIS Citation Index 中的被引频次:  
3  
被引频次合计:  
104  
ISSN:  
0926-5805  
eISSN:  
1872-7891  
入藏号:  
WOS:000928177900003  
Indexed Date:  
2023-03-10  
Institution Address:  
Southeast Univ, Sch Transportat, Dept Roadway Engn, Nanjing 211189, Peoples R China  
Institution Address:  
Southeast Univ, Natl Demonstrat Ctr Expt Rd & Traff Engn Educ, Nanjing 211189, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Artificial Intelligence, Nanjing 210044, Peoples R China  
Institution:  
Institution:  
Institution:

### 305. A novel dimensional variational prototypical network for industrial

摘要: A Dimensional Variational Prototypical Network (DVPN) is proposed to learn transferable knowledge from a largescale dataset containing sufficient samples of diverse faults, enabling few-shot diagnosis on new faults that are unseen in the dataset. The network includes a multiscale feature fusion module with shared weights to extract fault features, followed by a dimensional variational prototypical module that uses variational inference to determine metric scaling parameters. This adaptive approach accurately measures feature similarity between samples and fault prototypes. To enhance discriminability, a representation learning loss is employed, distinguishing between the least similar samples within the same class (hard positive samples) and the most similar samples across different classes (hard negative samples). The network combines representation learning and prototypical learning through the joint representation learning (JRL) module, acquiring both task-level and feature-level knowledge for a more discriminative metric space and improved classification accuracy on unseen faults. Experimental evaluations on datasets from the Tennessee Eastman process and a real-world polyester esterification process show that the proposed DVPN achieves high diagnostic performance and is comparable to state-of-the-art methods for few-shot fault diagnosis (FSFD).  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0166-3615  
eISSN:  
1872-6194  
入藏号:  
WOS:001284859700001  
Indexed Date:  
2024-08-13  
Institution Address:  
Donghua Univ, Engn Res Ctr Digitized Text & Apparel Technol, Minist Educ, Coll Informat Sci & Technol, Shanghai 201620, Peoples R China  
Institution:

### 306. Underwater object detection algorithm based on feature enhancement and

摘要: To solve the problems that the conventional object detector is hard to extract features and miss detec-tion of small objects when detecting underwater objects due to the noise of underwater environment and the scale change of objects, this paper designs a novel feature enhancement & progressive dynamic aggregation strategy, and proposes a new underwater object detector based on YOLOv5s. Firstly, a fea-ture enhancement gating module is designed to selectively suppress or enhance multi-level features and reduce the interference of underwater complex environment noise on feature fusion. Then, the adjacent feature fusion mechanism and dynamic fusion module are designed to dynamically learn fusion weights and perform multi-level feature fusion progressively, so as to suppress the conflict information in multi -scale feature fusion and prevent small objects from being submerged by the conflict information. At last, a spatial pyramid pool structure (FMSPP) based on the same size quickly mixed pool layer is proposed, which can make the network obtain stronger description ability of texture and contour features, reduce the parameters, and further improve the generalization ability and classification accuracy. The ablation experiments and multi-method comparison experiments on URPC and DUT-USEG data sets prove the effectiveness of the proposed strategy. Compared with the current mainstream detectors, our detector achieves obvious advantages in detection performance and efficiency.(c) 2023 Published by Elsevier Ltd.  
在 BIOSIS Citation Index 中的被引频次:  
4  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
19  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
20  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:000952571400001  
Indexed Date:  
2023-04-05  
Institution Address:  
Naval Univ Engn, Natl Key Lab Sci & Technol Vessel Integrated Power, Wuhan 430033, Peoples R China  
Institution:

### 307. Vehicle Detection Based on Adaptive Multimodal Feature Fusion and

摘要: Target detection is a critical task in interpreting aerial images. Small target detection, such as vehicles, is challenging. Different lighting conditions affect the accuracy of vehicle detection. For example, vehicles are difficult to distinguish from the background in red, green, blue (RGB) images under low illumination conditions. In contrast, under high-illumination conditions, the color and texture of vehicles are not significantly different in thermal infrared (TIR) images. To improve the accuracy of vehicle detection under various illumination conditions, we propose an adaptive multimodal feature fusion and cross-modal vehicle index (AFFCM) model for vehicle detection. Based on the single-stage object detection model, AFFCM uses RGB and TIR images. It comprises three parts: 1) the softpooling channel attention (SCA) mechanism calculates the cross-modal feature weights of the RGB and TIR features using a fully connected layer during global weighted pooling; 2) we design a multimodal adaptive feature fusion (MAFF) module based on the cross-modal feature weights derived from the SCA mechanism; the MAFF selects features with high weight, compresses redundant features with low weight, and performs adaptive fusion using a multiscale feature pyramid; and 3) a cross-modal vehicle index is established to extract the target area, suppress complex background information, and minimize false alarms in vehicle detection. The mean average precision (mAP) on the Drone Vehicle dataset is 14.44% and 5.02% higher than that obtained using only RGB or TIR images. The mAP is 2.63% higher than that of state-of-the-art methods that utilize RGB and TIR images.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
15  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
16  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001081620400014  
Indexed Date:  
2023-11-01  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Key Lab Computat Opt Imaging Technol, Beijing 100094, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 100049, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Optoelect, Beijing 100049, Peoples R China  
Institution:  
Institution:  
Institution:

### 308. MSFFAL: Few-Shot Object Detection via Multi-Scale Feature Fusion and

摘要: Few-shot object detection (FSOD) is proposed to solve the application problem of traditional detectors in scenarios lacking training samples. The meta-learning methods have attracted the researchers' attention for their excellent generalization performance. They usually select the same class of support features according to the query labels to weight the query features. However, the model cannot possess the ability of active identification only by using the same category support features, and feature selection causes difficulties in the testing process without labels. The single-scale feature of the model also leads to poor performance in small object detection. In addition, the hard samples in the support branch impact the backbone's representation of the support features, thus impacting the feature weighting process. To overcome these problems, we propose a multi-scale feature fusion and attentive learning (MSFFAL) framework for few-shot object detection. We first design the backbone with multi-scale feature fusion and channel attention mechanism to improve the model's detection accuracy on small objects and the representation of hard support samples. Based on this, we propose an attention loss to replace the feature weighting module. The loss allows the model to consistently represent the objects of the same category in the two branches and realizes the active recognition of the model. The model no longer depends on query labels to select features when testing, optimizing the model testing process. The experiments show that MSFFAL outperforms the state-of-the-art (SOTA) by 0.7-7.8% on the Pascal VOC and exhibits 1.61 times the result of the baseline model in MS COCO's small objects detection.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
1424-8220  
入藏号:  
WOS:000969003700001  
PubMed ID:  
37050671  
Indexed Date:  
2023-04-28  
Institution Address:  
Chinese Acad Sci, Shanghai Adv Res Inst, Shanghai 201210, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 100049, Peoples R China  
Institution Address:  
ShanghaiTech Univ, Sch Informat Sci & Technol, Shanghai 201210, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Rock & Soil Mech, State Key Lab Geomech & Geotech Engn, Wuhan 430071, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 309. Research on Multi-Scale Feature Fusion Network Algorithm Based on Brain

摘要: Gliomas have the highest mortality rate of all brain tumors. Correctly classifying the glioma risk period can help doctors make reasonable treatment plans and improve patients' survival rates. This paper proposes a hierarchical multi-scale attention feature fusion medical image classification network (HMAC-Net), which effectively combines global features and local features. The network framework consists of three parallel layers: The global feature extraction layer, the local feature extraction layer, and the multi-scale feature fusion layer. A linear sparse attention mechanism is designed in the global feature extraction layer to reduce information redundancy. In the local feature extraction layer, a bilateral local attention mechanism is introduced to improve the extraction of relevant information between adjacent slices. In the multi-scale feature fusion layer, a channel fusion block combining convolutional attention mechanism and residual inverse multi-layer perceptron is proposed to prevent gradient disappearance and network degradation and improve feature representation capability. The double-branch iterative multi-scale classification block is used to improve the classification performance. On the brain glioma risk grading dataset, the results of the ablation experiment and comparison experiment show that the proposed HMAC-Net has the best performance in both qualitative analysis of heat maps and quantitative analysis of evaluation indicators. On the dataset of skin cancer classification, the generalization experiment results show that the proposed HMAC-Net has a good generalization effect.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001268215500005  
Indexed Date:  
2024-07-20  
Institution Address:  
China Acad Engn Phys, Grad Sch, Beijing 100000, Peoples R China  
Institution Address:  
Beijing Inst Appl Phys & Computat Math, Dept Computat Math, Beijing 100000, Peoples R China  
Institution Address:  
Natl Key Lab Computat Phys, Beijing 100088, Peoples R China  
Institution Address:  
Shanghai Zhangjiang Inst Math, Shanghai 201203, Peoples R China  
Institution:  
Institution:  
Beijing Inst Appl Phys & Computat Math  
Institution:  
Natl Key Lab Computat Phys  
Institution:  
Shanghai Zhangjiang Inst Math

### 310. Dual encoding DDS-UNet liver tumour segmentation based on multi-scale

摘要: The fusion and utilization of multi-scale deep and shallow features are of great significance in liver tumour segmentation. This study proposes a dual encoding DDS-UNet liver tumour segmentation method based on multi-scale deep and shallow feature fusion, aiming to fully achieve the fusion and utilization of deep and shallow features and achieve accurate segmentation. The proposed method mainly consists of residual convolution module fusion residual convolution (FRC), dual encoding end fusion module dual encoding end fusion (DEF), and skip connection fusion module jump connection fusion module (JCF). In the residual convolution module, a layer by layer fused residual convolution is used instead of traditional convolution to achieve better training. In the dual encoding end fusion module, multi-scale feature fusion at the end provides more comprehensive contextual information, solves the loss of spatial geometric information. The skip connection fusion module reduces the interference of invalid features by changing the weights of spatial attention and channel attention on important features. The Dice coefficient, average intersection to union ratio, accuracy, recall, and accuracy indicators tested on the LiTS dataset were 90.37%, 90.16%, 93.78%, 94.91%, and 98.84%, respectively, which are superior to many advanced liver tumor segmentation methods.  
A dual encoding DDS-UNet liver tumour segmentation model based on multi-scale deep and shallow feature fusion is proposed to address the issue of inaccurate image segmentation caused by inconsistent lesion size, unclear boundary lines, and overly complex network in medical images.image  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1751-9659  
eISSN:  
1751-9667  
入藏号:  
WOS:001129810800001  
Indexed Date:  
2024-01-08  
Institution Address:  
Chongqing Normal Univ, Coll Comp & Informat Sci, Chongqing, Peoples R China  
Institution Address:  
Chongqing Normal Univ, Coll Comp & Informat Sci, Chongqing, Peoples R China  
Institution:  
Institution:

### 311. Effective image tampering localization with multi-scale ConvNeXt feature

摘要: With the widespread use of powerful image editing tools, image tampering becomes easy and realistic. Existing image forensic methods still face challenges of low generalization performance and robustness. In this letter, we propose an effective image tampering localization scheme based on ConvNeXt encoder and multi-scale Feature Fusion (ConvNeXtFF). Stacked ConvNeXt blocks are utilized as an encoder to capture hierarchical multi-scale features, which are then fused in decoder for locating tampered pixels accurately. Combined loss function and effective data augmentation strategies are adopted to further improve the model performance. Extensive experimental results show that both localization accuracy and robustness of the ConvNeXtFF scheme outperform other state-of-the-art ones. The source code is available at https://github.com/multimediaFor/ConvNeXtFF.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
17  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
17  
ISSN:  
1047-3203  
eISSN:  
1095-9076  
入藏号:  
WOS:001133634100001  
Indexed Date:  
2024-01-11  
Institution Address:  
Commun Univ China, Sch Comp & Cyber Sci, Beijing 100024, Peoples R China  
Institution Address:  
Commun Univ China, State Key Lab Media Convergence & Commun, Beijing 100024, Peoples R China  
Institution Address:  
Peoples Publ Secur Univ China, Beijing 100038, Peoples R China  
Institution:  
Institution:  
Institution:

### 312. AM-MSFF: A Pest Recognition Network Based on Attention Mechanism and

摘要: Traditional methods for pest recognition have certain limitations in addressing the challenges posed by diverse pest species, varying sizes, diverse morphologies, and complex field backgrounds, resulting in a lower recognition accuracy. To overcome these limitations, this paper proposes a novel pest recognition method based on attention mechanism and multi-scale feature fusion (AM-MSFF). By combining the advantages of attention mechanism and multi-scale feature fusion, this method significantly improves the accuracy of pest recognition. Firstly, we introduce the relation-aware global attention (RGA) module to adaptively adjust the feature weights of each position, thereby focusing more on the regions relevant to pests and reducing the background interference. Then, we propose the multi-scale feature fusion (MSFF) module to fuse feature maps from different scales, which better captures the subtle differences and the overall shape features in pest images. Moreover, we introduce generalized-mean pooling (GeMP) to more accurately extract feature information from pest images and better distinguish different pest categories. In terms of the loss function, this study proposes an improved focal loss (FL), known as balanced focal loss (BFL), as a replacement for cross-entropy loss. This improvement aims to address the common issue of class imbalance in pest datasets, thereby enhancing the recognition accuracy of pest identification models. To evaluate the performance of the AM-MSFF model, we conduct experiments on two publicly available pest datasets (IP102 and D0). Extensive experiments demonstrate that our proposed AM-MSFF outperforms most state-of-the-art methods. On the IP102 dataset, the accuracy reaches 72.64%, while on the D0 dataset, it reaches 99.05%.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
1  
eISSN:  
1099-4300  
入藏号:  
WOS:001232900600001  
PubMed ID:  
38785680  
Indexed Date:  
2024-06-03  
Institution Address:  
Xinjiang Univ, Sch Comp Sci & Technol, Urumqi 830017, Peoples R China  
Institution Address:  
Xinjiang Univ, Xinjiang Key Lab Multilingual Informat Technol, Urumqi 830017, Peoples R China  
Institution:  
Institution:

### 313. SAFFNet: Self-Attention-Based Feature Fusion Network for Remote Sensing

摘要: In real applications, it is necessary to classify new unseen classes that cannot be acquired in training datasets. To solve this problem, few-shot learning methods are usually adopted to recognize new categories with only a few (out-of-bag) labeled samples together with the known classes available in the (large-scale) training dataset. Unlike common scene classification images obtained by CCD (Charge-Coupled Device) cameras, remote sensing scene classification datasets tend to have plentiful texture features rather than shape features. Therefore, it is important to extract more valuable texture semantic features from a limited number of labeled input images. In this paper, a multi-scale feature fusion network for few-shot remote sensing scene classification is proposed by integrating a novel self-attention feature selection module, denoted as SAFFNet. Unlike a pyramidal feature hierarchy for object detection, the informative representations of the images with different receptive fields are automatically selected and re-weighted for feature fusion after refining network and global pooling operation for a few-shot remote sensing classification task. Here, the feature weighting value can be fine-tuned by the support set in the few-shot learning task. The proposed model is evaluated on three publicly available datasets for few shot remote sensing scene classification. Experimental results demonstrate the effectiveness of the proposed SAFFNet to improve the few-shot classification accuracy significantly compared to other few-shot methods and the typical multi-scale feature fusion network.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
42  
在 BIOSIS Citation Index 中的被引频次:  
3  
被引频次合计:  
42  
eISSN:  
2072-4292  
入藏号:  
WOS:000670980600001  
Indexed Date:  
2021-07-19  
Institution Address:  
Fudan Univ, Sch Comp Sci, Shanghai Key Lab Data Sci, 2005 SongHu Rd, Shanghai 200438, Peoples R China  
Institution Address:  
Zhongshan PoolNet Technol Co Ltd, Zhongshan Fudan Joint Innovat Ctr, 6 Xiangxing Rd, Zhongshan 528400, Peoples R China  
Institution:  
Institution:  
Zhongshan PoolNet Technol Co Ltd

### 314. Establishing effective learning bridge cross multi-scale feature maps

摘要: In the field of image object detection and semantic segmentation, improving the accuracy of object identification and segmentation is a primary goal. To achieve this, leveraging the potential of multi-scale information through feature map refinement and fusion has been widely recognized. However, existing feature fusion methods either design more complex feature pyramid networks, replace existing detectors, or incrementally introduce feature fusion modules, overlooking the effective approach of enhancing spatial information in deep feature maps. We propose a novel pluggable feature fusion paradigm termed 'Effective Learning Bridge'. Our research introduces an efficient and adaptive learning mechanism that builds learning bridges between feature maps at different scales within the feature pyramid, thereby enhancing the spatial information of objects in deep feature maps. This mechanism is specifically designed for multi-scale feature maps and can be seamlessly integrated into any network incorporating feature maps. By altering the model's backpropagation path, we successfully improve learning efficiency, which in turn enhances the accuracy of object detection and segmentation. Our proposed paradigm and method were extensively evaluated through experiments on SIMD, HRSID, and WHDLD datasets and benchmark models. The results unequivocally demonstrate the effectiveness of our approach in significantly improving the accuracy of object detection and semantic segmentation, as well as the overall learning efficiency of the model.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:001346504300001  
Indexed Date:  
2024-11-12  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Astronaut, Nanjing, Jiangsu, Peoples R China  
Institution Address:  
Zhejiang Univ, Coll Phys, Hangzhou, Zhejiang, Peoples R China  
Institution:  
Institution:

### 315. Real-Time Forestry Pest Detection Method Based on Enhanced Feature

摘要: The aim of this study is to address the real-time requirements of forestry pest detection and the problem of a low detection rate caused by anchor box redundancy of existing detection methods. This paper proposes a real-time forestry pest detection method based on theanchor-free method that can balance the detection rate and detection accuracy. Based on the TTFNet method, a mobile feature extraction network is introduced, and the effective feature weights are increased by one-dimensional convolution before feature output to suppress invalid features. For pest detection, data are mostly small-scale targets. An enhanced feature fusion method is proposed to introduce an asymmetric convolution module in multi-scale feature fusion to feature-enhance the feature maps extracted by the backbone network and connect across layers to improve the detection accuracy. To address the degradation of the anchor box position regression loss in the original method, DIOULoss is introduced to optimize the position regression loss function of the anchor box. Finally, data augmentation is performed on a relatively small number of samples in the dataset, the accuracy of the model is improved by 1.94%, the FPS is improved to 1.6 times of the original one, and the training time is slightly increased compared with the preinnovation model. Ablation experiments are designed to demonstrate the effectiveness of the proposed algorithm while being more conducive to deployment on edge devices.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2090-0147  
eISSN:  
2090-0155  
入藏号:  
WOS:001135121200001  
Indexed Date:  
2022-08-28  
Institution Address:  
Lanzhou Univ Technol, Sch Comp & Commun, Lanzhou, Gansu, Peoples R China  
Institution:

### 316. MSFYOLO: Feature fusion-based detection for small objects

摘要: At present, the effect of object detection algorithm in small object detection is very poor, mainly because the low-level network lacks semantic information and the characteristic information expressed by small object inspection data is very lack. In view of the above difficulties, this paper proposes a small object detection algorithm based on multi-scale feature fusion. By learning shallow features at the shallow level and deep features at the deep level, the proposed multi-scale feature learning scheme focuses on the fusion of concrete features and abstract features. It constructs object detector (MSFYOLO) based on multi-scale deep feature learning network and considers the relationship between a single object and local environment. Combining global information with local information, the feature pyramid is constructed by fusing different depth feature layers in the network. In addition, this paper also proposes a new feature extraction network (CourNet), through the way of feature visualization compared with the mainstream backbone network, the network can better express the small object feature information. The proposed algorithm is valuated on the MS COCO and achieved leading performance with 11.7% improvement in FPS, 17.0% improvement in AP, 81.0% improvement in ARS, and 23.3% reduction in computational FPLOs compared to YOLOv3. This study shows that the combination of global information and local information is helpful to detect the expression of small objects in different illumination. MSFYOLO uses CourNet as the backbone network, which has high efficiency and a good balance between accuracy and speed.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
25  
在中国科学引文数据库中的被引频次:  
6  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
32  
ISSN:  
1548-0992  
入藏号:  
WOS:000747446400015  
Indexed Date:  
2022-02-03  
Institution Address:  
Hebei Normal Univ Sci & Technol, Qinhuangdao 066000, Hebei, Peoples R China  
Institution Address:  
Hebei Univ Sci & Technol, Sch Informat Sci & Engn, Shijiazhuang 050018, Hebei, Peoples R China  
Institution:  
Institution:

### 317. Welding Defect Monitoring Based on Multi-Scale Feature Fusion of Molten

摘要: Real-time quality monitoring through molten pool images is a critical focus in researching high-quality, intelligent automated welding. However, challenges such as the dynamic nature of the molten pool, changes in camera perspective, and variations in pool shape make defect detection using single-frame images difficult. We propose a multi-scale fusion method for defect monitoring based on molten pool videos to address these issues. This method analyzes the temporal changes in light spots on the molten pool surface, transferring features between frames to capture dynamic behavior. Our approach employs multi-scale feature fusion using row and column convolutions along with a gated fusion module to accommodate variations in pool size and position, enabling the detection of light spot changes of different sizes and directions from coarse to fine. Additionally, incorporating mixed attention with row and column features enables the model to capture the characteristics of the molten pool more efficiently. Our method achieves an accuracy of 97.416% on a molten pool video dataset, with a processing time of 16 ms per sample. Experimental results on the UCF101-24 and JHMDB datasets also demonstrate the method's generalization capability.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001341566300001  
PubMed ID:  
39460041  
Indexed Date:  
2024-11-03  
Institution Address:  
Shandong Univ Sci & Technol, Coll lntelligent Equipment, Tai An 271019, Peoples R China  
Institution Address:  
Beijing Botsing Technol Co Ltd, Beijing 100176, Peoples R China  
Institution:  
Institution:  
Beijing Botsing Technol Co Ltd

### 318. Visual Inspection Method for Metal Rolls Based on Multi-Scale Spatial

摘要: Metal rolls in a non-ferrous-metal manufacturing workshop manifest the characteristics of symmetry, multiple scales and mutual covering, which poses great challenges for metal roll detection. To solve this problem, firstly, an efficient attention mechanism algorithm named ECLAM (efficient capture location attendant model) is proposed for capturing spatial position features efficiently, to obtain complete location information for metal rolls in a complex environment. ECLAM can improve the ability to extract the spatial features of backbone networks and reduce the influence of the non-critical background. In addition, in order to give feature maps a larger receptive field and improve the weight of location information in multi-scale feature maps, a nonlinear feature fusion module named LFFM (location feature fusion module) is used to fuse two adjacent feature images. Finally, a multi-scale object detection network named L-MSNet (location-based multi-scale object detection network) based on the combination of ECLAM and LFFM is proposed and used to accurately detect multi-scale metal rolls. In the experiments, multi-scale metal roll images are collected from an actual non-ferrous-metal manufacturing workshop. On this basis, a pixel-level image dataset is constructed. Comparative experiments show that, compared with other object detection methods, L-MSNet can detect multi-scale metal rolls more accurately. The average accuracy is improved by 2% to 5%, and the average accuracy of small and medium-sized objects is also significantly improved by 3% to 6%.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2073-8994  
入藏号:  
WOS:000834488300001  
Indexed Date:  
2022-08-10  
Institution Address:  
Cent South Univ, Sch Automat, Changsha 410083, Peoples R China  
Institution Address:  
Chinese Acad Sci, Shenyang Inst Automat, State Key Lab Robot, Shenyang 110016, Peoples R China  
Institution Address:  
China Nonferrous Met Proc Technol Co Ltd, Luoyang 410005, Peoples R China  
Institution Address:  
Hunan Normal Univ, Coll Engn & Design, Changsha 410081, Peoples R China  
Institution:  
Institution:  
Institution:  
China Nonferrous Met Proc Technol Co Ltd  
Institution:

### 319. YOLO-APDM: Improved YOLOv8 for Road Target Detection in Infrared Images

摘要: A new algorithm called YOLO-APDM is proposed to address low quality and multi-scale target detection issues in infrared road scenes. The method reconstructs the neck section of the algorithm using the multi-scale attentional feature fusion idea. Based on this reconstruction, the P2 detection layer is established, which optimizes network structure, enhances multi-scale feature fusion performance, and expands the detection network's capacity for multi-scale complicated targets. Replacing YOLOv8's C2f module with C2f-DCNv3 increases the network's ability to focus on the target region while lowering the amount of model parameters. The MSCA mechanism is added after the backbone's SPPF module to improve the model's detection performance by directing the network's detection resources to the major road target detection zone. Experimental results show that on the FLIR\_ADAS\_v2 dataset retaining eight main categories, using YOLO-APDM compared to YOLOv8n, mAP@0.5 and mAP@0.5:0.95 increased by 6.6% and 5.0%, respectively. On the M3FD dataset, mAP@0.5 and mAP@0.5 increased by 8.1% and 5.9%, respectively. The number of model parameters and model size were reduced by 8.6% and 4.8%, respectively. The design requirements of the high-precision detection of infrared road targets were achieved while considering the requirements of model complexity control.  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:001366527700001  
PubMed ID:  
39598973  
Indexed Date:  
2024-12-07  
Institution Address:  
Nanchang Univ, Sch Informat Engn, Nanchang 330019, Peoples R China  
Institution:

### 320. Small-Target Pedestrian-Detection Algorithm Based on Improved YOLOv4

摘要: Pedestrian detection is vital to applications in unmanned environment perception.Most existing pedestriandetection algorithms focus only on ordinary pedestrian targets and do not consider the low accuracy caused by the insufficient pedestrian feature information of small targets;furthermore,they do not offer favorable real-time performance when applied to embedded devices.Hence,a small-target pedestrian-detection algorithm,YOLOv4-DBF,is proposed herein. The conventional convolution is replaced with deeply separable convolution in the YOLOv4 algorithm,which reduces the number of parameters and the computation time of the model,as well as improves the detection speed and real-time performance of the algorithm.Additionally,the concurrent spatial and channel Squeeze & Excitation(scSE) attention module is introduced into the feature fusion component of the YOLOv4 backbone network to enhance the important channels and spatial features of the input pedestrian feature map as well as to enable the network to learn more meaningful feature information. The feature fusion component of the Feature Pyramid Network(FPN) in the YOLOv4 neck is improved to enhance the multiscale feature learning of the pedestrian target in the image,which improves the detection accuracy but increases the amount of computation.After training and verification based on the VOC07+12+COCO dataset,the results show that compared with the original YOLOv4 algorithm,YOLOv4-DBF increases the Average Precision(AP) by 4.16 percentage points and the speed by 27%.Finally,YOLOv4-DBF is accelerate deployed on the TX2 equipment of an unmanned vehicle for real-time testing,where the maximum speed reaches 23FPS.The algorithm proposed herein can effectively improve the accuracy and real-time performance of small-target pedestrian detection.  
摘要:  
行人检测在无人驾驶环境感知领域具有重要应用。现有行人检测算法多数只关注普通大小的行人目标,忽略了小目标行人特征信息过少的问题,从而造成检测精度低、应用于嵌入式设备中实时性不高等情况。针对该问题,提出一种小目标行人检测算法YOLOv4-DBF。引用深度可分离卷积代替YOLOv4算法中的传统卷积,以降低模型的参数量和计算量,提升检测速度和算法实时性。在YOLOv4骨干网络中的特征融合部分引入scSE注意力模块,对输入行人特征图的重要通道和空间特征进行增强,促使网络学习更有意义的特征信息。对YOLOv4颈部中特征金字塔网络的特征融合部分进行改进,在增加少量计算量的情况下增强对图像中行人目标的多尺度特征学习,从而提高检测精度。在VOC07+12+COCO数据集上进行训练和验证,结果表明,相比原YOLOv4算法, YOLOv4-DBF算法的AP值提高4.16个百分点,速度提升27%,将该算法加速部署在无人车中的TX2设备上进行实时测试,其检测速度达到23FPS,能够有效提高小目标行人检测的精度及实时性。  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
5  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1000-3428  
入藏号:  
CSCD:7409032  
Indexed Date:  
2023-06-10  
Institution Address:  
北京联合大学;;北京联合大学机器人学院, 北京市信息服务工程重点实验室;;, ;;, 北京;;北京 100101;;100101, 中国  
Institution Address:  
Beijing Union University;;College of Robotics,Beijing Union University, Beijing Key Laboratory of Information Service Engineering;;, ;;, Beijing;;Beijing 100101;;100101  
Institution:  
北京联合大学;;北京联合大学机器人学院  
Institution:  
Beijing Union University;;College of Robotics,Beijing Union University

### 321. Semantic Segmentation of Remote Sensing Images Based on Dual Attention

摘要: We propose a remote sensing image semantic segmentation model based on dual attention and multi-scale feature fusion to solve the problems of objects scale differences and missing small objects. This model uses ResNet50 in the coding part to extract features. First of all, the output features of each stage of ResNet50 are introduced into the pyramid pooling module, making full use of the multi-scale context information of the image to cope with the change of the object scales. Secondly, the dual attention is introduced in the final output features of ResNet50 to establish the semantic relationship between the spatial and channel dimensions, which enhances the ability of feature representation and improve the condition that small targets are difficult to segment. Finally, starting from the output features of the attention module, the features of all levels are gradually integrated to complete decoding to refine the target segmentation edge. The designed comparative experiments results show the effectiveness of the proposed method.  
会议名称:  
12th International Conference on Graphics and Image Processing (ICGIP)  
会议日期:  
NOV 13-15, 2020  
会议地点:  
Xian, PEOPLES R CHINA  
会议赞助商:  
Foshan Univ; Changan Univ; NW A & F Univ; Sichuan Univ; Ocean Univ China; Univ Portsmouth; China Soc Image & Graph; China Graph Soc; Xian Univ Technol  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0277-786X  
eISSN:  
1996-756X  
ISBN:  
978-1-5106-4278-2; 978-1-5106-4277-5  
入藏号:  
WOS:000700371000034  
Indexed Date:  
2021-10-24  
Institution Address:  
Jiangxi Normal Univ, Sch Comp Informat Engn, Nanchang 330022, Jiangxi, Peoples R China  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Informat & Commun Engn, Beijing 100876, Peoples R China  
Institution:  
Institution:

### 322. Multi-scale network toward real-world image denoising

摘要: Images are inevitably degraded when captured due to the effects of noise, and thus denoising is required. Previous methods remove real-world noise, while also causing issues with over-smoothing image details and loss of edge information. To solve these issues, a multi-scale image denoising network (MSIDNet) is proposed in this paper. We design a residual attention block (RAB) to encode and decode the context well, while introducing a selective kernel feature fusion module to fuse multi-scale features and obtain rich contextual information from low-resolutions to restore more details. A feature extraction block (FEB) is designed to fully extract local and global features then fusion, which obtains rich feature information. Extensive experiments on four real-world image datasets demonstrate that our method has excellent generalization and achieves advanced denoising performance on both peak signal-to-noise ratio and structural similarity. MSIDNet preserves more edge details and improves the over-smoothing issue to enhance the visual effect of denoised images.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
1868-8071  
eISSN:  
1868-808X  
入藏号:  
WOS:000879697100001  
Indexed Date:  
2022-11-18  
Institution Address:  
Yunnan Univ, Sch Informat Sci & Engn, Kunming 650504, Yunnan, Peoples R China  
Institution:

### 323. Research on Vehicle Detection Method Based on Improved YOLOX-s

摘要: A improved vehicle detection model based on multi-scale feature fusion of YOLOX network is proposed to solve the problem of missing and false detection of small vehicle targets. Ghost-cross stage partial(CSP) based on the depth separable convolution is designed to replace part of cross stage partial in network to speed up the speed of detection. The max pooling mode of model is improved to Softpool mode, and coordinate attention mechanism is introduced to enhance the feature expression of target to be detected and to optimize the problem of target missing detection. Focal Loss is selected as the confidence loss function of model to increase the weight of inaccurate classification samples and improve the prediction ability of the model for small targets. The experimental results show that the average accuracy of the improved algorithm is improved to 74.96%, and the speed is up to 73 frames per second, which can better meet the requirements of real-time vehicle target detection.  
摘要:  
为缓解车辆小目标漏检及误检问题,提出一种基于YOLOX网络的多尺度特征融合的改进车辆检测模型。设计基于深度可分离卷积的Ghost-CSP(cross stage partial),替换网络的部分跨阶段局部结构,加快检测速度;将模型的最大池化方式改进为Softpool方式,并引入坐标注意力机制,增强待检测目标的特征表达,优化目标漏检问题;选用Focal Loss作为模型置信度损失函数以增加分类不准确样本的权重,提高模型对小目标的预测能力。实验结果表明:改进算法平均准确率提高到74.96%,速度达到73帧/s,在满足实时性要求下可以更好地完成车辆目标检测要求。  
在中国科学引文数据库中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1004-731X  
入藏号:  
CSCD:7651450  
Indexed Date:  
2024-04-12  
Institution Address:  
江苏理工学院电气信息工程学院, 常州, 213001  
Institution Address:  
College of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou, 213001  
Institution:  
江苏理工学院电气信息工程学院  
Institution:  
College of Electrical and Information Engineering, Jiangsu University of Technology

### 324. Pedestrian detection based on channel feature fusion and enhanced

摘要: At present, pedestrian detection is widely applied to autonomous driving and intelligent transportation and robots, etc. But the balance between accuracy and speed is still not reached. In complex background with high pedestrian density and serious occlusion, missing detection or false detection may occur by pedestrian detection models based on center and scale prediction (CSP). An improved pedestrian detection method based on channel feature fusion and enhanced semantic segmentation is presented. A feature fusion module based on squeeze and excitation is proposed in feature extraction. Multi-scale feature maps are fused to obtain faster detection speed and higher detection accuracy. An enhanced semantic segmentation module is presented in detection head to solve missing detection for long-distance pedestrians. CIOU (Complete Intersection Over Union) loss function is used to improve the confidence levels of pedestrians. Experiments on different networks, scales of feature fusion and detection methods are carried out to verify the performance of proposed approach. The experimental results show that the proposed model can detect pedestrians with high accuracy in occluded, dense and long-distance scenes. The detection speed can be accelerated while keeping low missed detection rate and less computational cost. It is shown that the approach can achieve high accuracy and robustness especially in complex background.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:001101209500006  
Indexed Date:  
2023-12-03  
Institution Address:  
Hubei Univ Technol, Sch Comp Sci, Wuhan 430068, Hubei, Peoples R China  
Institution:

### 325. Weakly Supervised Building Change Detection Based on DeepCut and

摘要: Building change detection for remote sensing images is crucial for urban development and illegal building management. Semantic segmentation networks (SSNs) are widely used for building change detection due to their multiscale feature extraction capabilities but require large amounts of pixel-level labeled data, which is time-consuming to annotate. To reduce annotation time, pixel-level pseudo-labels using graph cuts (GrabCuts) for interactive foreground extraction can be generated by only selecting bounding boxes for changed building regions. However, GrabCut faces challenges in remote sensing images due to complex environments. In addition, changes in seasons and lighting can alter building appearances, leading to "pseudo-changes" in detection results. Surrounding objects close to buildings can also interfere with feature fusion, making it difficult to accurately extract changed building features. To address these issues, this article proposes DeepCut and the temporal invariant change detection network (TICDN). DeepCut enhances GrabCut by incorporating deep features that provide higher level semantic information, resulting in more accurate segmentation of building parts. TICDN introduces adversarial learning to reduce the impact of pseudo-changes caused by seasonal and lighting variations, ensuring robust building features. It also features a correlation-based feature fusion module (CFM) for feature fusion, which adaptively calculates the correlation between prechange and postchange features. This module effectively eliminates the influence of irrelevant objects, such as roads and trees, and accurately extracts changed building features. Extensive comparative experiments and ablation studies on the WHU and LEVIR-CD datasets demonstrate the superior performance and broad applicability of these methods in building change detection. The code is available at https://github.com/YuchunHuang/Weakly-Supervised-Change-Detection.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001294099800009  
Indexed Date:  
2024-08-24  
Institution Address:  
Wuhan Univ, Sch Remote Sensing & Informat Engn, Wuhan 430072, Peoples R China  
Institution Address:  
Univ Penn, Sch Engn & Appl Sci, Dept Elect & Syst Engn, Philadelphia, PA 19104 USA  
Institution Address:  
Alibaba Grp, Hangzhou, Peoples R China  
Institution:  
Institution:  
Institution:

### 326. Towards High Accuracy Pedestrian Detection on Edge GPUs

摘要: Despite the rapid development of pedestrian detection algorithms, the balance between detection accuracy and efficiency is still far from being achieved due to edge GPUs (low computing power) limiting the parameters of the model. To address this issue, we propose the YOLOv4-TP-Tiny based on the YOLOv4 model, which mainly includes two modules, two-dimensional attention (TA) and pedestrian-based feature extraction (PFM). First, we integrate the TA mechanism into the backbone network, which increases the attention of the network to the visible area of pedestrians and improves the accuracy of pedestrian detection. Then, the PFM is used to replace the original spatial pyramid pooling (SPP) structure in the YOLOv4 to obtain the YOLOv4-TP algorithm, which can adapt to different sizes of people to obtain higher detection accuracy. To maintain detection speed, we replaced the normal convolution with a ghost network with a TA mechanism, resulting in more feature maps with fewer parameters. We constructed a one-way multi-scale feature fusion structure to replace the down-sampling process, thereby reducing network parameters to obtain the YOLOv4-TP-Tiny model. The experimental results show that the YOLOv4-TP-tiny has 58.3% AP and 31 FPS in the winder person pedestrian dataset. With the same hardware conditions and dataset, the AP of the YOLOv4-tiny is 55.9%, and the FPS is 29.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
15  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
15  
eISSN:  
1424-8220  
入藏号:  
WOS:000845244000001  
PubMed ID:  
36015741  
Indexed Date:  
2022-09-03  
Institution Address:  
Anhui Univ Sci & Technol, Sch Comp Sci & Engn, Huainan 232001, Peoples R China  
Institution Address:  
Tongji Univ, Coll Elect & Informat Engn, Shanghai 200092, Peoples R China  
Institution:  
Institution:

### 327. Airport small object detection based on feature enhancement

摘要: Video object detection is essential for airport surface surveillance, but the objects on the scene are mostly small objects with low resolution, they have no obvious feature information. Due to the scale differences of the objects and the fixed receptive field on the feature maps, detectors cannot model multi-scale context information and cover all objects. In addition, although the video detection algorithm can be used as a method to solve the problem of small object detection, the temporal feature fusion method of current video detection is very dependent on the quality of a single feature map. Therefore, this paper aims to enhance the features of small objects of a single image. First, an attentional multi-scale feature fusion enhancement (A-MSFFE) network is built on the memory-enhanced global-local aggregation (MEGA) to supplement semantic and spatial information of small objects. Then, a context feature enhancement (CFE) module is designed for obtaining different receptive fields through different dilated convolutions. Meanwhile, a video detection dataset about the airport is established. Finally, the experimental results show that the proposed method can improve the detection accuracies of small objects and outperform other state-of-the-art video object detection algorithms in self-built airport dataset.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1751-9659  
eISSN:  
1751-9667  
入藏号:  
WOS:000723369700001  
Indexed Date:  
2021-12-03  
Institution Address:  
Sichuan Univ, Sch Aeronaut & Astronaut, Chengdu 610065, Peoples R China  
Institution Address:  
Sichuan Univ, Coll Comp Sci, Chengdu, Peoples R China  
Institution:  
Institution:

### 328. A Lightweight YOLO Object Detection Algorithm Based on Bidirectional

摘要: This paper proposes a lightweight YOLO object detection algorithm based on bidirectional multi-scale feature enhancement. The problem is that the original YOLOv5 algorithm does not make full use of the relationship between the feature layers, resulting in the loss of target semantic information and a large number of parameters. First, a bidirectional multi-scale feature-enhanced weighted fusion backbone network is constructed to extract target features repeatedly. It enhances the fusion ability of shallow detail features and high-level semantic information to capture richer multi-scale semantic information. Second, the NCA attention module is built and integrated into the feature fusion network to enhance the critical characteristics of the target region. Finally, the Ghost module is used instead of the convolutional blocks in the original network to lighten the model while reducing the network complexity and training difficulty. Experimental results show that the improved YOLOv5 algorithm achieves 78.8% mAP@0.5 for the PASCAL VOC2012 dataset, which is 1.5% higher than the original algorithm, at 62.5 FPS. The number of parameters is also reduced by 43.6%. The mAP@0.5 on the self-made metal foreign object dataset reached 98.4%, at 58.8 FPS, which can meet the requirements of end-device deployment and real-time detection.  
In this paper, a bi-directional multi-scale feature-enhanced weighted fusion backbone is designed to enhance the fusion capability of shallow features and advanced features. The NCA attention module is designed and embedded into the feature fusion network to enhance the key features in the target region. The Ghost module is used to reduce the network complexity and training difficulty. image  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2513-0390  
入藏号:  
WOS:001182456100001  
Indexed Date:  
2024-03-20  
Institution Address:  
Henan Polytech Univ, Sch Elect Engn & Automat, Jiaozuo 454000, Henan, Peoples R China  
Institution Address:  
Henan Intelligent Equipment, Int Joint Lab Direct Drive & Control, Zhengzhou 454000, Henan, Peoples R China  
Institution Address:  
Muroran Inst Technol, Coll Informat & Syst, Muroran, Hokkaido 0508585, Japan  
Institution:  
Institution:  
Henan Intelligent Equipment  
Institution:

### 329. Insulator Defect Degree Detection Based on Multi-scale Feature Fusion

摘要: Insulator defects with different degrees have similar features and less pixel information, resulting in poor detection effect, therefore, an insulator defect degree detection network based on multi-scale feature fusion(MFFD3Net) is proposed. The network uses reconstructed ResNeSt50 to improve the feature extraction ability in insulator defect dataset. A multi-scale feature fusion module based on deconvolution is designed, which enriches the expression ability of different size feature maps and improves the detection performance of different scale targets. At the same time, the receptive field block(RFB) is added after the shallow feature maps of the input detection module to ensure more insulator defect information to enter the effective receptive field, which has an impact on the final feature map and improves the detection accuracy of insulator defects in different degrees. The mAP of MFFD3Net on insulator defect degree dataset reaches 85.02%, the detection accuracy of small targets such as slight breakage and slight flashover is 78.37% and 79.98%, which can complete the identification and location of insulator defects in different degrees. Thus, the MFFD3Net proposed in this paper is of great significance for improving the fault warning of power system and ensuring the safe and stable operation of power grid.  
摘要:  
针对绝缘子不同程度缺陷特征相似、像素信息少、不同程度缺陷检测效果不佳的问题,提出了一种基于多尺度特征融合的绝缘子缺陷程度检测网络(multi-scale feature fusion defect degree detection network, MFFD3Net)。该网络采用重构的ResNeSt50架构提高了对绝缘子缺陷程度数据集的特征提取能力。设计了基于反卷积的多尺度特征融合模块,丰富了不同尺寸特征图的表达能力,提高了对不同尺度目标的检测性能。同时,在输入检测模块的浅层特征图后增加多感受野的特征提取模块(receptive field block,RFB),使得更多绝缘子缺陷信息进入有效感受野,对最终特征图产生影响,提升不同程度绝缘子缺陷的检测精度。MFFD3Net在绝缘子缺陷程度数据集上的全类平均精度达到85.02%,其中绝缘子轻微破损与绝缘子轻微闪络小目标的检测精度分别为78.37%、79.98%,能够完成不同程度绝缘子缺陷的识别与定位。因此,该文提出的MFFD3Net对于完善电力系统故障预警、保障电网安全稳定运行具有重要意义。  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1003-6520  
入藏号:  
CSCD:7741134  
Indexed Date:  
2024-08-23  
Institution Address:  
(徐州)中国矿业大学电气工程学院, 徐州, 江苏 221000, 中国  
Institution Address:  
国网江苏省电力有限公司徐州供电分公司, 徐州, 江苏 221000, 中国  
Institution Address:  
School of Electrical Engineering, China University of Mining and Technology (Xuzhou), Xuzhou, Jiangsu 221000, China  
Institution Address:  
Xuzhou Power Supply Company, State Grid Jiangsu Electric Power Compang, Xuzhou, Jiangsu 221000, China  
Institution:  
(徐州)中国矿业大学电气工程学院  
Institution:  
国网江苏省电力有限公司徐州供电分公司  
Institution:  
School of Electrical Engineering, China University of Mining and Technology (Xuzhou)  
Institution:  
Xuzhou Power Supply Company, State Grid Jiangsu Electric Power Compang

### 330. A safety helmet-wearing detection method based on cross-layer connection

摘要: Given the current safety helmet detection methods, the feature information of the small-scale safety helmet will be lost after the network model is convolved many times, resulting in the problem of missing detection of the safety helmet. To this end, an improved target detection algorithm of YOLOv5 is used to detect the wearing of safety helmets. Firstly, a new small-scale detection layer is added to the head of the network for multi-scale feature fusion, thereby increasing the receptive field area of the feature map to improve the model's recognition of small targets. Secondly, a cross-layer connection is designed between the feature extraction network and the feature fusion network to enhance the fine-grained features of the target in the shallow layer of the network. Thirdly, a coordinate attention (CA) module is added to the cross-layer connection to capture the global information of the image and improve the localization ability of the target. Finally, the Normalized Wasserstein Distance (NWD) is used to measure the similarity between bounding boxes, replacing the intersection over union (IoU) method. The experimental results show that the improved model achieves 95.09% of the mAP value for safety helmet-wearing detection, which has a good effect on the recognition of small-sized safety helmets of different degrees in the construction work scene.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001199294800001  
Indexed Date:  
2024-04-14  
Institution Address:  
Xihua Univ, Sch Elect Engn & Elect Informat, Chengdu 610039, Peoples R China  
Institution:

### 331. A Video Target Re-Recognition Method Based on Adaptive Attention

摘要: Within the realm of computer vision, the task of re-identifying targets across multiple video frames has emerged as a pivotal challenge, particularly in domains like video surveillance, smart transportation systems, and pedestrian flow analytics. Conventional re-identification techniques often grapple with constraints stemming from varying camera perspectives, inconsistent lighting conditions, and prevalent occlusions. Addressing these challenges, this research introduces MVF-Re, a sophisticated re-identification approach that synergizes adaptive attention mechanisms with multi-scale feature fusion. Initially, we architect a deep attention-enhanced feature pyramid network, a pioneering framework that dynamically tailors itself to video frame content, thereby capturing intricate target details. Subsequently, we incorporate a multi-input Siamese network, ensuring the derivation of consistent and resilient feature sets across diverse contexts. To augment feature distinctiveness, we conceptualize a context-sensitive dynamic attention mechanism, adept at judiciously allocating weights to individual video frames. Culminating our approach, we deploy an innovative multi-scale feature fusion methodology, offering a holistic and robust target representation. Empirical evaluations on multiple benchmark datasets underscore the superior performance of our methodology, underscoring its proficiency in multi-frame target re-identification.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001150379400001  
Indexed Date:  
2024-02-04  
Institution Address:  
Guangzhou Xinhua Univ, Dongguan 523133, Peoples R China  
Institution:  
Guangzhou Xinhua Univ

### 332. Using Attention Mechanism and Improved YOLOX Vehicle Taillight Detection

摘要: Taillight detection is of great significance and value in predicting driving intention of vehicle ahead in assisted driving and unmanned driving systems. Aiming at the problem that the vehicle taillight detection model YOLOX is not able to detect small targets and the channel information does not contain location information, a taillight detection algorithm based on the attention mechanism and improved YOLOX is proposed. Firstly, a micro-scale detection layer is added to further multi-scale feature fusion to extract more taillight feature information. Secondly, the mobile network attention mechanism is inserted into the proposed improved feature fusion network to extract taillight position information. Finally, the sensing field area at the output end of feature extraction layer is enlarged to improve the detection accuracy of small targets such as taillights. The taillight data set established in natural environment was used, and the number of samples was 1,962 images. The training set, verification set and test set were randomly divided in 8:1:1 ratio to verify the improved network. Experimental results show that the improved model has a 91.71% average detection accuracy (mAP) for taillight images, which is 3.48% higher than the standard YOLOX algorithm.  
会议名称:  
15th International Conference on Machine Vision (ICMV)  
会议日期:  
NOV 18-20, 2022  
会议地点:  
Rome, ITALY  
会议赞助商:  
Zhonghui Exhibit Grp; Univ Elect Sci & Technol China; Sapienza Univ Rome  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0277-786X  
eISSN:  
1996-756X  
ISBN:  
978-1-5106-6618-4; 978-1-5106-6619-1  
入藏号:  
WOS:001012785800002  
Indexed Date:  
2023-07-23  
Institution Address:  
Nanjing Univ Sci & Technol, Nanjing, Peoples R China  
Institution:

### 333. Global-Feature-Fusion and Multiscale Network for Low-Frequency

摘要: Full waveform inversion (FWI) is currently the most accurate technique for obtaining the properties of subsurface media. The absence of low frequencies in the observed data caused cycle-skipping phenomenon and poor initial model which affect the convergence of FWI. We propose a global-feature-fusion and multiscale network (GM-Net) in a way of supervised learning to compensate for the absent low-frequency components in the observed data trace by trace. The difficulty of extrapolating frequency is to achieve smoothness and continuity when changing from high-frequency signals to low-frequency signals, which is visually shown in the reduction and movement of the sidelobes in high-frequency signals and the overall oscillation of the signals is slowed down. For achieving better extrapolation, the encoder-decoder architecture with multiscale feature extraction is designed as the backbone of the network. For avoiding the loss of information, we propose to perform 1/2 downsampling on the original input signal separately based on the odd and even time samples and then concatenate them along the channel dimension. Since 1-D seismic data are a type of time-series signal and the wavelengths of low frequencies are long, we pay more attention to the relevance of contextual information. Thus, dilated convolution layers, gridding convolution blocks, and nonlocal attention blocks are used to enlarge the receptive field in both time and channel dimensions to extract and fuse global features. Numerical tests on both synthetic data and different types of field marine data demonstrate the feasibility and generalization of our method.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001294606500006  
Indexed Date:  
2024-08-28  
Institution Address:  
Minist Educ, Key Lab Modern Power Syst Simulat & Control & Rene, Jilin 132012, Peoples R China  
Institution Address:  
Northeast Elect Power Univ, Coll Elect Engn, Dept Commun Engn, Jilin 132012, Peoples R China  
Institution Address:  
Jilin Univ, Coll Instrumentat & Elect Engn, Changchun 130026, Peoples R China  
Institution Address:  
Jilin Univ, Coll Geoexplorat Sci & Technol, Changchun 130026, Peoples R China  
Institution:  
Minist Educ  
Institution:  
Institution:  
Institution:

### 334. A Lightweight and Dynamic Feature Aggregation Method for Cotton Field

摘要: Weed detection is closely related to agricultural production, but often faces the problems of leaf shading and limited computational resources. Therefore, this study proposes an improved weed detection algorithm based on YOLOv8. Firstly, the Dilated Feature Integration Block is designed to improve the feature extraction in the backbone network by introducing large kernel convolution and multi-scale dilation convolution, which utilizes information from different scales and levels. Secondly, to solve the problem of a large number of parameters in the feature fusion process of the Path Aggregation Feature Pyramid Network, a new feature fusion architecture multi-scale feature interaction network is designed, which achieves the high-level semantic information to guide the low-level semantic information through the attention mechanism. Finally, we propose a Dynamic Feature Aggregation Head to solve the problem that the YOLOv8 detection head cannot dynamically focus on important features. Comprehensive experiments on two publicly accessible datasets show that the proposed model outperforms the benchmark model, with mAP50 and mAP75 improving by 4.7% and 5.0%, and 5.3% and 3.3%, respectively, whereas the number of model parameters is only 6.62 M. This study illustrates the utility potential of the algorithm for weed detection in cotton fields, marking a significant advancement of artificial intelligence in agriculture.  
Web of Science 核心合集中的 "被引频次":  
3  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
2079-9292  
入藏号:  
WOS:001245366900001  
Indexed Date:  
2024-06-18  
Institution Address:  
Xinjiang Univ, Sch Informat Sci & Engn, Urumqi 830017, Peoples R China  
Institution Address:  
Xinjiang Univ, Xinjiang Key Lab Multilingual Informat Technol, Urumqi 830017, Peoples R China  
Institution Address:  
Xinjiang Teachers Coll, Sch Informat Sci & Technol, Urumqi 830043, Peoples R China  
Institution:  
Institution:  
Institution:  
Xinjiang Teachers Coll

### 335. A Remote Sensing Target Detection Model Based on Lightweight Feature

摘要: Remote sensing image (RSI) target detection methods based on traditional multiscale feature fusion (MSFF) have achieved great success. However, the traditional MSFF method significantly increases the computational cost during model training and inference, and the simple fusion operation may lead to the semantic confusion of the feature map, which cannot realize the refined extraction of features by the model. In order to reduce the computational effort associated with the MSFF operation and to enable the features in the feature map to present an accurate, fine-grained distribution, we propose a single-stage detection model (RS-YOLO). Our main additions to RS-YOLO are a computationally smaller and faster Quick and Small E-ELEN (QS-E-ELEN) module and a feature refinement extraction (FRE) module. In the QS-E-ELEN module, We utilize QSBlock, jump-join, and convolution operations to fuse features on different scales and reduce the computational effort of the model by exploiting the similarity of the RSI feature map channels. In order for the model to better utilize the enhanced features, we designed the FRE module to make the location of the enhanced features more accurate and fine. By conducting experiments on the popular NWPU-VHR- 10 and SSDD datasets, we derive results showing that RS-YOLO outperforms most mainstream models in terms of the tradeoff between accuracy and speed. Specifically, in terms of accuracy, it improves 1.6% and 1.7% compared to the current state-of-the-art models, respectively. At the same time, RS-YOLO reduces the number of parameters and computational effort.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001225641400011  
Indexed Date:  
2024-06-07  
Institution Address:  
Nanyang Inst Technol, Sch Comp & Software, Nanyang 473004, Peoples R China  
Institution Address:  
Henan Univ, Sch Comp & Informat Engn, Kaifeng 475001, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Chongqing Engn Res Ctr Spatial Big Data Intelligen, Chongqing 400065, Peoples R China  
Institution:  
Institution:  
Institution:

### 336. LUD-YOLO: A novel lightweight object detection network for unmanned

摘要: Autonomous execution of tasks by unmanned aerial vehicles (UAVs) relies heavily on object detection. However, object detection in most images presents challenges such as complex backgrounds, small targets, and obstructions. Additionally, the limited computing speed and memory of the UAV processor affects the accuracy of conventional object detection algorithms. This paper proposes LUD-You Only Look Once (YOLO), a small and lightweight object detection algorithm for UAVs based on YOLOv8. The proposed algorithm introduces a new multiscale feature fusion mode that solves the degradation in feature propagation and interaction through the introduction of upsampling in the feature pyramid network and the progressive feature pyramid network. The application of the dynamic sparse attention mechanism in the Cf2 module achieves flexible computing allocation and content awareness. Furthermore, the proposed model is optimized to be sparse and lightweight, making it possible to deploy on UAV edge devices. Finally, the effectiveness and superiority of LUD-YOLO were verified on the VisDrone2019 and UAVDT datasets. The results of ablation and comparison experiments show that compared with the original algorithm, LUDY-N and LUDY-S have shown excellent performance in various evaluation indexes, indicating that the proposed improvement strategies make the model have better robustness and generalization. Moreover, compared with multiple other popular competitors, the proposed improvement strategies enable LUD-YOLO to have the best overall performance, providing an effective solution for UAVs object detection while balancing model size and detection accuracy.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
0020-0255  
eISSN:  
1872-6291  
入藏号:  
WOS:001301189800001  
Indexed Date:  
2024-09-04  
Institution Address:  
Huazhong Univ Sci & Technol, State Key Lab Intelligent Mfg Equipment & Technol, Wuhan 430074, Peoples R China  
Institution Address:  
Guizhou Univ Finance & Econ, Coll Big Data Stat, Guiyang 550025, Peoples R China  
Institution Address:  
Guizhou Univ, Key Lab Adv Mfg Technol, Minist Educ, Guiyang 550025, Guizhou, Peoples R China  
Institution Address:  
Natl Def Univ, Turkish Naval Acad, Dept Ind Engn, TR-34942 Istanbul, Turkiye  
Institution Address:  
Imperial Coll London, Royal Sch Mines, South Kensington Campus, London SW7 2AZ, England  
Institution Address:  
Western Caspian Univ, Dept Informat Technol, Baku 1001, Azerbaijan  
Institution Address:  
Zhejiang Univ, Coll Informat Sci & Elect Engn, Hangzhou 310058, Peoples R China  
Institution Address:  
Lebanese Amer Univ, Dept Comp Sci & Math, Beirut, Lebanon  
Institution Address:  
Middle East Univ, MEU Res Unit, Amman 11831, Jordan  
Institution:  
Institution:  
Institution:  
Institution:  
Natl Def Univ  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 337. Dense and Small Object Detection in UAV-Vision Based on a Global-Local

摘要: Unmanned aerial vehicles (UAVs) have been widely used in postdisaster search and rescue operations, object tracking, and other tasks. Therefore, the autonomous perception of UAVs based on computer vision has become a research hotspot in recent years. However, UAV images include dense objects, small objects, and arbitrary object directions, which bring about significant challenges to existing object detection methods. To alleviate these issues, we propose a global-local feature enhanced network (GLF-Net). Considering the difficulty of processing UAV images with complex scenes and dense objects, we designed a backbone based on an involution and self-attention that can extract effective features from complex objects. A multiscale feature fusion module is also proposed to address the presence of numerous small objects in UAV images through multiscale object detection and feature fusion. To accurately detect rotated objects, a rotated regional proposal network was designed based on the midpoint offset representation, which can apply a rotated box to determine the real direction and contour of an object. GLF-Net achieves a state-of-the-art detection accuracy [86.52% mean average precision (mAP)] on our created rotated object detection UAV (RO-UAV) dataset, while achieving 96.95% and 97% mAP on the public datasets high resolution ship collections 2016 (HRSC2016) and the University of Chinese Academy of Sciences High Resolution Aerial Object Detection Dataset (UCAS-AOD), respectively. The experimental results demonstrate that our method achieves a high detection accuracy and generalization, which can meet the practical requirements of UAVs under various complex scenarios.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
33  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
33  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000843270700007  
Indexed Date:  
2022-08-30  
Institution Address:  
China Univ Min & Technol, Sch Mech Elect & Informat Engn, Beijing 100083, Peoples R China  
Institution Address:  
State Key Lab Coal Min & Clean Utilizat, Beijing 100083, Peoples R China  
Institution Address:  
Minist Emergency Management, Key Lab Intelligent Min & Robot, Beijing 100083, Peoples R China  
Institution:  
Institution:  
State Key Lab Coal Min & Clean Utilizat  
Institution:  
Minist Emergency Management

### 338. Research on Digital Media Intelligent Art Creation Based on the Fusion

摘要: With the rapid development of computer science, technologies such as virtual reality and semantic feature fusion have been greatly developed in recent years. In order to construct a digital media intelligent art creation model, this paper first designs a digital media intelligent creation framework based on virtual reality and maintains the original pattern and color distribution of natural images as accurately as possible. This method first performs automatic color segmentation and clustering on the image to extract key lines; then, re-colors the different segmented regions; and finally completes the detail enhancement. For the optimization of color matching, the algorithm integrates the contrast of the color block, and uses the differential evolution strategy to optimize the color configuration scheme. In addition, this paper also proposes a generative adversarial network-based on semantic feature fusion for digital media intelligent art creation. The network introduces multi-scale feature fusion by embedding the residual block feature pyramid structure and directly generates the final fine image by adaptively fusing these features. Only one discriminator can generate a realistic image of 256px. 256px. The experimental results have shown that this paper makes full use of virtual reality and semantic feature fusion technology, and obtains better results and better performance of intelligent art creation results on digital media than existing methods.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1016-2364  
入藏号:  
WOS:001040536900001  
Indexed Date:  
2023-08-12  
Institution Address:  
Xinyang Agr & Forestry Univ, Coll Planning & Design, Xinyang 46400, Peoples R China  
Institution Address:  
Zhengzhou Inst Aeronaut Ind Management, Sch Humanities & Law, Zhengzhou 450046, Peoples R China  
Institution:  
Xinyang Agr & Forestry Univ  
Institution:

### 339. A Defect Detection Method Based on BC-YOLO for Transmission Line

摘要: Vibration dampers and insulators are important components of transmission lines, and it is therefore important for the normal operation of transmission lines to detect defects in these components in a timely manner. In this paper, we provide an automatic detection method for component defects through patrolling inspection by an unmanned aerial vehicle (UAV). We constructed a dataset of vibration dampers and insulators (DVDI) on transmission lines in images obtained by the UAV. It is difficult to detect defects in vibration dampers and insulators from UAV images, as these components and their defective parts are very small parts of the images, and the components vary greatly in terms of their shape and color and are easily confused with the background. In view of this, we use the end-to-end coordinate attention and bidirectional feature pyramid network "you only look once" (BC-YOLO) to detect component defects. To make the network focus on the features of vibration dampers and insulators rather than the complex backgrounds, we added the coordinate attention (CA) module to YOLOv5. CA encodes each channel separately along the vertical and horizontal directions, which allows the attention module to simultaneously capture remote spatial interactions with precise location information and helps the network locate targets of interest more accurately. In the multiscale feature fusion stage, different input features have different resolutions, and their contributions to the fused output features are usually unequal. However, PANet treats each input feature equally and simply sums them up without distinction. In this paper, we replace the original PANet feature fusion framework in YOLOv5 with a bidirectional feature pyramid network (BiFPN). BiFPN introduces learnable weights to learn the importance of different features, which can make the network focus more on the feature mapping that contributes more to the output features. To verify the effectiveness of our method, we conducted a test in DVDI, and its mAP@0.5 reached 89.1%, a value 2.7% higher than for YOLOv5.  
在中国科学引文数据库中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
39  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
42  
eISSN:  
2072-4292  
入藏号:  
WOS:000873517100001  
Indexed Date:  
2022-11-02  
Institution Address:  
Anhui Univ, Sch Elect & Informat Engn, Hefei 230601, Peoples R China  
Institution Address:  
Chinese Acad Sci, Hefei Inst Phys Sci, Hefei 230031, Peoples R China  
Institution:  
Institution:

### 340. MEF-UNet: An end-to-end ultrasound image segmentation algorithm based on

摘要: Ultrasound image segmentation is a challenging task due to the complexity of lesion types, fuzzy boundaries, and low-contrast images along with the presence of noises and artifacts. To address these issues, we propose an end-to-end multi-scale feature extraction and fusion network (MEF-UNet) for the automatic segmentation of ultrasound images. Specifically, we first design a selective feature extraction encoder, including detail extraction stage and structure extraction stage, to precisely capture the edge details and overall shape features of the lesions. In order to enhance the representation capacity of contextual information, we develop a context information storage module in the skip-connection section, responsible for integrating information from adjacent two-layer feature maps. In addition, we design a multi-scale feature fusion module in the decoder section to merge feature maps with different scales. Experimental results indicate that our MEF-UNet can significantly improve the segmentation results in both quantitative analysis and visual effects.  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
15  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
15  
ISSN:  
0895-6111  
eISSN:  
1879-0771  
入藏号:  
WOS:001217545500001  
PubMed ID:  
38513396  
Indexed Date:  
2024-05-16  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Math & Stat, Nanjing 210044, Jiangsu, Peoples R China  
Institution Address:  
Zhejiang Univ, Sch Math Sci, Hangzhou 310027, Zhejiang, Peoples R China  
Institution Address:  
Chinese Univ Hong Kong, Dept Math, Shatin, Hong Kong, Peoples R China  
Institution:  
Institution:  
Institution:

### 341. MSFF: A Multi-Scale Feature Fusion Network for Surface Defect Detection

摘要: This paper presents an improved YOLOv3 network, named MSFF-YOLOv3, for precisely detecting variable surface defects of aluminum profiles in practice. First, we introduce a larger prediction scale to provide detailed information for small defect detection; second, we design an efficient attention-guided block to extract more features of defects with less overhead; third, we design a bottom-up pyramid and integrate it with the existing feature pyramid network to construct a twin-tower structure to improve the circulation and fusion of features of different layers. In addition, we employ the K-median algorithm for anchor clustering to speed up the network reasoning. Experimental results showed that the mean average precision of the proposed network MSFF-YOLOv3 is higher than all conventional networks for surface defect detection of aluminum profiles. Moreover, the number of frames processed per second for our proposed MSFF-YOLOv3 could meet real-time requirements.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
0916-8532  
eISSN:  
1745-1361  
入藏号:  
WOS:000852731400019  
Indexed Date:  
2022-09-18  
Institution Address:  
Shaanxi Univ Sci & Technol, Xian, Peoples R China  
Institution Address:  
Aichi Prefectural Univ, Fac Informat Sci & Technol, Nagakute, Aichi 4801198, Japan  
Institution:  
Institution:  
Aichi Prefectural Univ

### 342. Person re-identification method based on global and local relation

摘要: A person re-identification method based on feature fusion and multi-scale information was proposed to solve the problem of low accuracy of person re-identification due to the large difference of human image background and similar global appearance of human body. Firstly,the global feature map of human body image is extracted by ResNet50. Secondly,the branch structure is designed. In the first branch,the spatial transformation network is used to align the global feature images adaptively,and the local feature images are obtained by horizontal segmentation of the global feature images. The correlation between the global feature and each local feature is mined by fusing the global feature and each local feature separately. The second branch adds four convolution layers of different scales to extract multi-scale features from global images. Finally,in the reasoning stage,the features of the first branch and the second branch are connected in series as the comparative features of person. Experiments on the Market-1501 and DukeMTMC datasets show that the proposed method has better performance than the AlignedReID and EA-NET feature alignment and local feature extraction methods. In the Market-1501 dataset,mAP and Rank-1 reach 86.77% and 94.83%,respectively.  
摘要:  
针对因行人图像背景差异大、人体外观相似导致的行人再识别准确率低的问题,提出了一种利用特征融合与多尺度信息的行人重识别方法。首先,通过ResNet50\_IBN提取人体图像全局特征图。其次,设计分支结构,第1分支利用空间变换网络对全局特征图进行自适应的空间特征对齐,水平切分全局特征图得到局部特征,采用全局特征与每个局部特征分别融合的方式来挖掘特征之间的关联关系。第2分支增加了4种不同尺度的卷积层提取全局图像的多尺度特征。最后,在推理阶段将第1分支和第2分支的特征进行通道维度的串联,作为行人的对比特征。通过在Market-1501、DukeMTMC数据集上的实验表明,所提方法与AlignedReID和EA-Net等特征对齐和局部特征提取方法相比具备更强的性能,在Market-1501上,mAP和Rank-1分别达到了86.77%和94.83%。  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
4  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1000-7105  
入藏号:  
CSCD:7267286  
Indexed Date:  
2022-10-21  
Institution Address:  
昆明理工大学信息工程与自动化学院, 昆明, 云南 650504, 中国  
Institution Address:  
广东省科学院智能制造研究所, 广东省现代控制技术重点实验室, 广州, 广东 510095, 中国  
Institution Address:  
School of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, Yunnan 650504, China  
Institution Address:  
Institute of Intelligent Manufacturing, Guangdong Academy of Sciences, Guangzhou, Guangdong 510095, China  
Institution:  
昆明理工大学信息工程与自动化学院  
Institution:  
广东省科学院智能制造研究所  
Institution:  
School of Information Engineering and Automation, Kunming University of Science and Technology  
Institution:  
Institute of Intelligent Manufacturing, Guangdong Academy of Sciences

### 343. Semantic segmentation of substation tools using an improved ICNet

摘要: In the field of substation operation and maintenance, real-time detection and precise segmentation of tools play an important role in maintaining the safe operation of the power grid and guiding operators to work safely. To improve the accuracy and real-time performance of semantic segmentation of substation operation and maintenance tools, we have proposed an improved, lightweight, real-time, semantic segmentation network based on an efficient image cascade network architecture (ICNet). The network uses multiscale branches and cascaded feature fusion units to extract rich multilevel features. We designed a semantic segmentation and purification module to deal with redundant and conflicting information in multiscale feature fusion. A lightweight backbone network was used in the feature extraction stage at different resolutions, and a recursive gated convolution was used in the upsampling stage to achieve high-order spatial interactions, thereby improving segmentation accuracy. Due to the lack of a substation tool semantic segmentation data set, we constructed one. Training and testing on the data set showed that the proposed model improved the accuracy of tool detection while ensuring real-time performance. Compared with the currently popular semantic segmentation network, it had better performance in real-time and accuracy, and provided a new semantic segmentation method for embedded platforms.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2688-1594  
入藏号:  
WOS:001318059700006  
Indexed Date:  
2024-10-11  
Institution Address:  
Beijing Informat Sci & Technol Univ, Sch Instrument Sci & Optoelect Engn, Beijing 100096, Peoples R China  
Institution Address:  
Beijing Kedong Elect Control Syst Co Ltd, Beijing 100192, Peoples R China  
Institution Address:  
NARI Grp Corp, State Grid Elect Power Res Inst, Nanjing 211106, Peoples R China  
Institution:  
Institution:  
Beijing Kedong Elect Control Syst Co Ltd  
Institution:

### 344. Lightweight Apple-Leaf Pathological Recognition Based on Multiscale

摘要: The occurrence of apple leaf diseases has a significant impact on apple quality and yield.Disease monitoring is therefore an important measure to ensure the healthy development of the apple industry.Based on the ResNet structure,a lightweight disease recognition model based on multiscale feature fusion is proposed.First,the feature fusion mechanism is used to extract and fuse the high-dimensional and low-dimensional features of the network,strengthen the transmission of semantic information between convolution layers,and enhance the ability to distinguish subtle lesions.Next,multiscale depth separable convolution is added to extract disease features of different scales by using multi-scale convolution kernel structure,which improves the richness of features and restricts the parameters of the model.Finally,a dataset containing five kinds of apple leaf diseases is used to verify the effectiveness of the proposed method.The experimental results show that the recognition accuracy of the model is 98.05%,and that the number and calculation of the model network are only 4.02 MB and 0.92 GB,respectively.Compared with other models,it also has advantages,and can provide a new scheme for the accurate identification of diseases and pests in agricultural automation.  
摘要:  
苹果叶部病害的发生极大地影响了苹果的品质和产量,对病害的监测是确保苹果产业健康发展的重要措施。在ResNet结构基础上,提出了一种基于多尺度特征融合的轻量化病害识别模型。首先,采用特征融合机制,提取并融合网络高低维特征,加强卷积层之间语义信息的传递,增强识别细微病斑的能力。其次,加入多尺度深度可分离卷积,利用多尺度卷积核结构对不同尺度的病害特征进行提取,提高特征的丰富度,同时约束模型的参数量。最后,为验证所提模型的有效性,采用一个包含5种苹果叶部病害数据集进行了实验。实验结果表明,该模型取得了98.05%的识别准确率,模型参数量和计算量仅为4.02 MB和0.92 GB,与其他模型相比同样具有优势,可为农业自动化精准识别病虫害提供新的方案。  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7404571  
Indexed Date:  
2023-06-02  
Institution Address:  
贵州大学大数据与信息工程学院;;半导体功率器件可靠性教育部工程研究中心, ;;半导体功率器件可靠性教育部工程研究中心, 贵阳;;贵阳, 贵州;;贵州 550025;;550025, 中国  
Institution Address:  
贵州大学机械工程学院, 贵阳, 贵州 550025, 中国  
Institution Address:  
贵州大学计算机科学与技术学院, 贵阳, 贵州 550025, 中国  
Institution Address:  
College of Big Data and Information Engineering,Guizhou University;;Semiconductor Power Device Reliability Engineering Research Center of the Ministry of Education, ;;Semiconductor Power Device Reliability Engineering Research Center of the Ministry of Education, Guiyang;;Guiyang, Guizhou;;Guizhou 550025;;550025  
Institution Address:  
School of Mechanical Engineering,Guizhou University, Guiyang, Guizhou 550025, China  
Institution Address:  
School of Computer Science and Technology,Guizhou University, Guiyang, Guizhou 550025, China  
Institution:  
贵州大学大数据与信息工程学院;;半导体功率器件可靠性教育部工程研究中心  
Institution:  
贵州大学机械工程学院  
Institution:  
贵州大学计算机科学与技术学院  
Institution:  
College of Big Data and Information Engineering,Guizhou University;;Semiconductor Power Device Reliability Engineering Research Center of the Ministry of Education  
Institution:  
School of Mechanical Engineering,Guizhou University  
Institution:  
School of Computer Science and Technology,Guizhou University

### 345. Research on an Intelligent Identification Method for Wind Turbine Blade

摘要: To address challenges in the detection of wind turbine blade damage images, characterized by complex backgrounds and multiscale feature distribution, we propose a method based on an enhanced YOLOV8 model. Our approach focuses on three key aspects: First, we enhance the extraction of small target features by integrating the CBAM attention mechanism into the backbone network. Second, the feature fusion process is refined using the Weighted Bidirectional Feature Pyramid Network (BiFPN) to replace the path aggregation network (PANet). This modification prioritizes small target features within the deep features and facilitates the fusion of multiscale features. Lastly, we improve the loss function from CIoU to EIoU, enhancing sensitivity to small targets and the perturbation resistance of bounding boxes, thereby reducing the gap between computed predictions and real values. Experimental results demonstrate that compared with the YOLOV8 model, the CBAM-BiFPN-YOLOV8 model exhibits improvements of 1.6%, 1.0%, 1.4%, and 1.1% in precision rate, recall rate, mAP@0.5, and mAP@0.5:.95, respectively. This enhanced model achieves substantial performance improvements comprehensively, demonstrating the feasibility and effectiveness of our proposed enhancements at a lower computational cost.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
14  
eISSN:  
2227-9717  
入藏号:  
WOS:001150789200001  
Indexed Date:  
2024-02-06  
Institution Address:  
Inner Mongolia Univ Sci & Technol, Sch Mech Engn, Baotou 014010, Peoples R China  
Institution Address:  
Inner Mongolia Key Lab Intelligent Diag & Control, Baotou 014010, Peoples R China  
Institution Address:  
Inner Mongolia Agr Univ, Coll Mech & Elect Engn, Hohhot 010018, Peoples R China  
Institution Address:  
Inner Mongolia Engn Res Ctr Intelligent Equipment, Hohhot 010018, Peoples R China  
Institution:  
Institution:  
Inner Mongolia Key Lab Intelligent Diag & Control  
Institution:  
Institution:  
Inner Mongolia Engn Res Ctr Intelligent Equipment

### 346. Building extraction method based on MFF-Deeplabv3+ network for

摘要: Automatic extraction of building information from high-resolution remote sensing images is of great significance in the fields of environmental monitoring, earthquake mitigation, and land use, making it a research hotspot in the field of high-resolution remote sensing applications. In order to improve the accuracy of building extraction from high-resolution remote sensing images, a building extraction method based on MFF-Deeplabv3+(multiscale feature fusion-Deeplabv3+) network for high-resolution remote sensing images is proposed in this paper. First, the multi-scale feature enhancement module is designed to enable the network to capture more scale context information; then, the feature fusion module is designed to effectively fuse deep features with shallow features to reduce the loss of detail information; finally, the attention mechanism module is introduced to select accurate features adaptively. In the comparison experiments of the Inria building dataset, MFFDeeplabv3+ achieved the highest accuracy in PA, MPA, FWIoU, and MIoU metrics with 95.75%, 91.22%, 92.12%, and 85.01%, respectively, while the generalization experiments of the WHU building dataset achieved good results. The results show that this method extracts building information from high-resolution remote sensing images with high accuracy and strong generalization.  
摘要:  
为提升高分辨率遥感影像中建筑物提取的精度,提出一种基于MFF-Deeplabv3+ (multiscale feature fusion-Deeplabv3+)网络的高分辨率遥感影像建筑物提取方法。首先,设计多尺度特征增强模块,使网络能够捕获更多尺度的上下文信息;然后,设计特征融合模块,有效融合深层特征与浅层特征,减少细节信息的丢失;最后,引入注意力机制模块,自适应地选择准确特征。在Inria建筑物数据集的对比实验中,MFF-Deeplabv3+在PA、MPA、FWIoU、MIoU指标中取得最高精度,分别为95.75% 、91.22% 、92.12%和85.01%,同时在WHU建筑物数据集的泛化实验中取得不错的结果。结果表明,本方法在高分辨率遥感影像中提取建筑物信息精度较高,且具有较好的泛化性。  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2095-6134  
入藏号:  
CSCD:7799471  
Indexed Date:  
2024-11-08  
Institution Address:  
中国科学院空天信息创新研究院;;中国科学院大学电子电气与通信工程学院, ;;, ;;, 北京;;北京 100094;;100049, 中国  
Institution Address:  
中国科学院空天信息创新研究院, 北京 100094, 中国  
Institution Address:  
中国科学院空天信息创新研究院;;中国科学院计算机网络信息中心;;中国科学院大学计算机科学与技术学院, ;;;;, ;;;;, 北京;;北京;;北京 100094;;100083;;100049, 中国  
Institution Address:  
Aerospace Information Research Institute,Chinese Academy of Sciences;;School of Electronic,Electrical and Communication Engineering,University of Chinese Academy of Sciences, ;;, ;;, Beijing;;Beijing 100094;;100049  
Institution Address:  
Aerospace Information Research Institute,Chinese Academy of Sciences, Beijing 100094, China  
Institution Address:  
Aerospace Information Research Institute,Chinese Academy of Sciences;;Computer Network Information Center,Chinese Academy of Sciences;;School of Computer Science and Technology,University of Chinese Academy of Sciences, ;;;;, ;;;;, Beijing;;Beijing;;Beijing 100094;;100083;;100049  
Institution:  
中国科学院空天信息创新研究院;;中国科学院大学电子电气与通信工程学院  
Institution:  
中国科学院空天信息创新研究院  
Institution:  
中国科学院空天信息创新研究院;;中国科学院计算机网络信息中心;;中国科学院大学计算机科学与技术学院  
Institution:  
Aerospace Information Research Institute,Chinese Academy of Sciences;;School of Electronic,Electrical and Communication Engineering,University of Chinese Academy of Sciences  
Institution:  
Aerospace Information Research Institute,Chinese Academy of Sciences  
Institution:  
Aerospace Information Research Institute,Chinese Academy of Sciences;;Computer Network Information Center,Chinese Academy of Sciences;;School of Computer Science and Technology,University of Chinese Academy of Sciences

### 347. Secure and Reliable Indoor Localization Based on Multitask Collaborative

摘要: Accurate and reliable indoor location estimate is crucial for many Internet-of-Things (IoT) applications in the era of smart buildings. However, the positioning accuracy and security of the existing positioning works cannot meet the demands in the large-scale smart buildings scenarios covering multiple multifloor buildings. Therefore, in this article, we focus on the reliable and accurate localization under multibuilding and multifloor environments. We propose two novel designs, including a two-step reliable feature selector and a multitask collaborative positioning model. First, we design a two-step reliable feature selector based on an access point (AP) confidence model and manifold learning, to help select the most representative and reliable fingerprint features. Second, we propose a multitask cooperative positioning model, which consists of a multiscale feature fusion module to adaptively fuse multiscale features and a multitask joint learning module to effectively constrain the cumulative error of multiscale position. Finally, based on the above two, we propose a reliable multibuilding and multifloor localization method (RMBMFL), which can achieve accurate and reliable location estimates with low computational complexity in a smart building complex. We did real-world experiments in a 20 000 m(2) site that covers three multistory buildings to evaluate the performance of the proposed RMBMFL. The experimental results show that RMBMFL achieves a building identification accuracy and a floor identification accuracy of 99%, and a room-level indoor localization with an average positioning error within 2 m, and outperforms state-of-the-art solutions.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
16  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
17  
ISSN:  
2327-4662  
入藏号:  
WOS:000879049400027  
Indexed Date:  
2022-11-18  
Institution Address:  
Hunan Univ, Coll Comp Sci & Elect Engn, Changsha 410082, Peoples R China  
Institution Address:  
Hunan Univ, Hunan Xiangjiang Artificial Intelligence Acad, Changsha 410082, Peoples R China  
Institution Address:  
Univ Technol Sydney, Sch Elect & Data Engn, Ultimo, NSW 2007, Australia  
Institution:  
Institution:  
Institution:

### 348. UcUNet: A lightweight and precise medical image segmentation network

摘要: In recent years, precise medical image segmentation methods based on the encoder-decoder structure have attracted much attention, but there are still some limitations. They mainly manifest in two aspects: one is that the network structure becomes increasingly complex in pursuit of segmentation ac-curacy, and the other is the insufficient multiscale information fusion ability between the encoder and decoder. To address these issues, we designed a novel lightweight precise medical image segmentation network called UcUNet, which has not only a very large receptive field and multiscale information fusion ability but also a low parameter count. Specifically, we first designed an efficient U-shaped convolution module that can increase the network depth with fewer parameters and effectively filter out invalid features while fusing shallow and deep features. Furthermore, we ingeniously introduced large-kernel convolution into this module, which significantly improved the network's receptive field with very few parameters, providing new ideas for the use of large-kernel convolution in lightweight models. In addition, we designed a self-adjusting multiscale feature fusion module based on large-kernel convolution to replace the original simple skip connections, effectively enhancing the network's multiscale information extraction and fusion ability. Finally, we conducted experiments on multiple datasets, and the results showed that our method can achieve precise medical image segmentation with a low parameter count.& COPY; 2023 Elsevier B.V. All rights reserved.  
在 BIOSIS Citation Index 中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
13  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
14  
ISSN:  
0950-7051  
eISSN:  
1872-7409  
入藏号:  
WOS:001063106600001  
Indexed Date:  
2023-09-27  
Institution Address:  
Southwest Univ Sci & Technol, Sch Informat Engn, Mianyang 621010, Peoples R China  
Institution Address:  
Southwest Univ Sci & Technol, Tianfu Inst Res & Innovat, Mianyang 621010, Peoples R China  
Institution Address:  
Mianyang Cent Hosp, Radiol Dept, Mianyang 621000, Peoples R China  
Institution:  
Institution:  
Institution:  
Mianyang Cent Hosp

### 349. Occluded Face Recognition Network Based on Deep Feature Suppression

摘要: Facial recognition technology is a key technology for verifying personal evidence in public security investigations. Although existing algorithms can achieve high recognition accuracy in unobstructed face recognition, effective facial features are lost when a face is occluded, resulting in a significant decrease in recognition accuracy. Hence, an occluded face recognition network based on deep feature suppression is proposed to address these issues. The network adaptively generates feature masks based on occluded faces, suppresses features damaged by occlusion in deep feature maps through feature masks, and uses the suppressed features to complete face recognition. To improve the discrimination of suppressed features, a twin network structure is used in the training phase to measure and learn the depth features of the occluded and corresponding unobstructed faces. Simultaneously, to fully utilize different levels of feature information, a Feature Pyramid Network (FPN) and an adaptive feature fusion module are constructed to extract multiscale feature information from faces. The feature layers containing more feature information are assigned greater fusion weights, thereby enhancing the representation abilities of the features. The experimental results show that the proposed method has good robustness, with accuracy rates of 99.50% and 98.42% for the LFW and LFW mask occlusions in the dataset, respectively, and 100%, 100%, 99.86%, and 99.02% for the four experimental settings in the AR dataset, respectively, surpassing those of current mainstream algorithms.  
摘要:  
人脸识别技术是公安侦查中人证核验的关键技术之一。尽管现有算法在无遮挡人脸识别上都能达到较高的识别精度,但当人脸被遮挡时,使得有效的人脸特征丢失,导致识别精度大幅下降。针对上述问题,提出一种基于深度特征抑制的遮挡人脸识别网络,通过遮挡人脸自适应地生成特征掩码,利用特征掩码抑制深层特征图中因遮挡损坏的特征,最后根据抑制后的特征完成人脸识别。为了提升抑制后特征的辨别力,在训练环节通过孪生网络结构将遮挡人脸与对应无遮挡人脸的深度特征进行度量学习。同时利用不同层次的特征信息,构建特征金字塔网络(FPN)和自适应特征融合模块对人脸的多尺度特征信息进行提取,对其中包含特征信息较多的特征层赋予更大的融合权重,从而增强特征的表征能力。实验结果表明,该方法具有较好的鲁棒性,其中在LFW数据集和LFW口罩遮挡数据集上的准确率分别达到了99.50%和98.42%,在AR数据集4个实验设置上的准确率分别达到了100%、100%、99.86%和99.02%,优于目前的主流算法。  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1000-3428  
入藏号:  
CSCD:7768607  
Indexed Date:  
2024-09-27  
Institution Address:  
西安邮电大学通信与信息工程学院, 西安, 陕西 710121, 中国  
Institution Address:  
中国科学院深圳先进技术研究院, 深圳, 广东 518055, 中国  
Institution Address:  
School of Communication and Information Engineering,Xi'an University of Posts and Telecommunications, Xi'an, Shaanxi 710121, China  
Institution Address:  
Shenzhen Institutes of Advanced Technology,Chinese Academy of Sciences, Shenzhen, Guangdong 518055, China  
Institution:  
西安邮电大学通信与信息工程学院  
Institution:  
中国科学院深圳先进技术研究院  
Institution:  
School of Communication and Information Engineering,Xi'an University of Posts and Telecommunications  
Institution:  
Shenzhen Institutes of Advanced Technology,Chinese Academy of Sciences

### 350. A terrain segmentation network for navigable areas with global strip

摘要: Accurate segmentation of safe navigable areas is crucial for scene parsing in autonomous driving systems. However, existing segmentation methods often fail to fully leverage the complementary nature of multiscale features in complex environments, resulting in inadequate information extraction. To overcome this limitation, a Terrain Segmentation Network (TSNet) is proposed for navigable area segmentation, which introduces reliable global semantic information as fusion guide while dynamically exploiting the complementary relationships between multiscale features for progressively guided fusion. Specifically, TSNet is composed of the Global Strip Reliability Evaluation Module (GSREM) and the Dynamic Fusion Module (DFM). The GSREM, which includes the Global Strip Unit (GSU) and the Reliability-Evaluation Unit (REU), is designed to capture the global geometric information of obstacles with long-range contextual characteristics. The information from GSREM serves as prior knowledge to guide the feature fusion. Additionally, we propose the DFM, comprising the Attention Fusion Unit (AFU) and the Contribution Distribution Unit (CDU), to explore the complementary relationships among multiscale feature interactions and to obtain comprehensive scene information. Extensive experiments on diverse wild datasets demonstrate that TSNet outperforms the state-of-the-art methods in accurately identifying navigable areas. The code for TSNet will be available at https://github.com/lv881314/ 2TSNet.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001385189000001  
Indexed Date:  
2025-01-03  
Institution Address:  
Shenzhen Univ, Guangdong Key Lab Intelligent Informat Proc, Shenzhen 518060, Guangdong, Peoples R China  
Institution Address:  
Shenzhen Univ, Shenzhen Key Lab Adv Machine Learning & Applicat, Shenzhen 518060, Guangdong, Peoples R China  
Institution Address:  
Shenzhen Univ, Coll Elect & Informat Engn, Shenzhen 518060, Guangdong, Peoples R China  
Institution Address:  
Jiangxi Agr Univ, Sch Elect & Informat Engn, Nanchang 330045, Jiangxi, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 351. Shadow detection via multi-scale feature fusion and unsupervised domain

摘要: Shadow detection is significant for scene understanding. As a common scenario, soft shadows have more ambiguous boundaries than hard shadows. However, they are rarely present in the available benchmarks since annotating for them is time-consuming and needs expert help. This paper discusses how to transfer the shadow detection capability from available shadow data to soft shadow data and proposes a novel shadow detection framework (MUSD) based on multi-scale feature fusion and unsupervised domain adaptation. Firstly, we set the existing labeled shadow dataset (i.e., SBU) as the source domain and collect an unlabeled soft shadow dataset (SSD) as the target domain to formulate an unsupervised domain adaptation problem. Next, we design an efficient shadow detection network based on the double attention module and multi-scale feature fusion. Then, we use the global-local feature alignment strategy to align the task-related feature distributions between the source and target domains. This allows us to obtain a robust model and achieve domain adaptation effectively. Extensive experimental results show that our method can detect soft shadows more accurately than existing state-of-the-art methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
被引频次合计:  
11  
ISSN:  
1047-3203  
eISSN:  
1095-9076  
入藏号:  
WOS:000849699300002  
Indexed Date:  
2022-09-10  
Institution Address:  
Hangzhou Dianzi Univ, Sch Comp Sci & Technol, Key Lab Complex Syst Modeling & Simulat, Hangzhou 310018, Peoples R China  
Institution:

### 352. Unbiased feature position alignment for human pose estimation

摘要: Multi-scale feature fusion is a commonly-used module in existing deep-learning models, and feature misalignment occurs in the process of feature fusion. The spatial misalignment hinders the learning of semantic representation with multi-scale levels, but which has not received much attention. This misalignment problem is caused by the feature position shift after using the convolution and interpolation operation in feature fusion. To solve the misalignment problem, this paper formulates the shift error mathematically and proposes a plug-and-play unbiased feature position alignment strategy to align convolution with interpolation. As a model-agnostic approach, unbiased feature position alignment can boost the performance of different models without introducing extra parameters. Furthermore, the unbiased feature position alignment is applied to build an unbiased human pose estimation method. Experimental results have demonstrated the effectiveness of the proposed unbiased pose model in comparison to the state-of-the-arts, especially in the low-resolution field. The codes are shared at https:// github.com/WangChen100/Unbiased-Feature-Position-Alignment-for-Human-Pose-Estimation.(c) 2023 Elsevier B.V. All rights reserved.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000977896900001  
Indexed Date:  
2023-05-15  
Institution Address:  
Hong Kong Polytech Univ, Hong Kong, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Nanjing 210003, Peoples R China  
Institution:  
Institution:

### 353. Substation instrumentation target detection based on multi-scale feature

摘要: With the promotion of smart grid construction work, the use of high-precision and high-efficiency substation inspection robot has become the development trend of substation inspection. A multi-scale feature fusion meter target detection algorithm is proposed to address the problems of low efficiency and susceptibility to surrounding environmental factors by the traditional manual meter reading method. Kinecct is used to acquire color images of substation meters with different backgrounds, light intensities, and angles to build a substation meter dataset. Based on the complementarity and correlation of multi-scale features, an SSD target detection model with multi-scale feature fusion is established, and the performance of the algorithm is tested on the constructed dataset, and comparative experiments are conducted to verify the effectiveness of the algorithm for target detection accuracy improvement.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
12  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
12  
ISSN:  
1532-0626  
eISSN:  
1532-0634  
入藏号:  
WOS:000823708500001  
Indexed Date:  
2022-07-20  
Institution Address:  
Wuhan Univ Sci & Technol, Minist Educ, Key Lab Met Equipment & Control Technol, Wuhan 430081, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Res Ctr Biomimet Robot & Intelligent Measurement, Wuhan 430081, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Coll Comp Sci & Technol, Wuhan 430081, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Hubei Prov Key Lab Intelligent Informat Proc & Re, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Hubei Key Lab Mech Transmiss & Mfg Engn, Wuhan, Peoples R China  
Institution Address:  
Three Gorges Univ, Hubei Key Lab Hydroelect Machinery Design & Maint, Yichang, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 354. Dangerous Goods Detection Based on Multi-Scale Feature Fusion in

摘要: Existing target detection algorithms have low accuracy in detecting smaller-sized dangerous goods in X-ray security inspection images. Therefore, a multi-scale feature fusion detection network called MFFNet (Multi-scale Feature Fusion Network) is proposed, which is based on the SSD detection model and uses a deeper feature extraction network, namely ResNet-101. The high-level semantic rich features of the network are merged with the low-level edge detailed features through the jump connection method, and contextual information is added for the detection of small-scale dangerous goods, which can effectively improve the identification and positioning accuracy of small scale targets. The new feature layer obtained by fusion and the SSD extended convolution layer are sent into detection together. Experimental results show that MFFNet can greatly improve the detection accuracy of dangerous goods in X-ray security inspection images, especially those of smaller sizes, while maintaining a relatively fast detection speed to meet the requirements of modern security inspection.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
ISSN:  
1006-4125  
入藏号:  
WOS:000686485400014  
Indexed Date:  
2021-09-10  
Institution Address:  
Civil Aviat Univ China, Tianjin Key Lab Intelligent Signal & Image Proc, Tianjin 300300, Peoples R China  
Institution:

### 355. Urban traffic tiny object detection via attention and multi-scale

摘要: The unmanned aerial vehicle (UAV) city patrol is of great significance in ensuring the safety of residents' lives and properties, as well as maintaining the normal operation of the city. However, the detection of UAV images faces challenges such as numerous small-scale objects, complex backgrounds, and high requirements for detection speed. In response to these issues, we introduce a Real-time Small Object Detection network in UAV-vision (RTS-Net), tailored for UAV patrols. Initially, we introduce a multiscale feature fusion module (MFFM) designed to augment the expressiveness of features across scales, thereby enhancing the detection of smaller objects. Subsequently, leveraging attention mechanisms, we present the coordinated attention detection module (CADM), which bolsters the detection model's ability to accurately segregate objects from the background in expansive, complex scenarios. Lastly, a lightweight real-time feature extraction module (RFEM) is crafted to diminish model computational complexity and boost inference speed. On the UAV road patrol image dataset we constructed, our proposed method attains a detection accuracy of 89.9%\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\%$$\end{document} mAP, breaking previous records. It surpasses all prevailing detection methods, particularly for small-scale objects. Simultaneously, it achieves an inference speed of 163.9 FPS. The experimental results show that RTS-Net can satisfy the accurate and efficient detection of ground objects by various different UAV platforms in different complex scenarios.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
2045-2322  
入藏号:  
WOS:001306841300032  
PubMed ID:  
39232041  
Indexed Date:  
2024-09-18  
Institution Address:  
Acad Mil Sci, Inst Syst Engn, Beijing 100000, Peoples R China  
Institution:  
Acad Mil Sci

### 356. Synthetic aperture radar target recognition via deep attention

摘要: Owing to the inherent characteristics of synthetic aperture radar (SAR) imaging system based on the coherent imaging modality, SAR images are inevitably corrupted by speckle noise, thereby affecting SAR target recognition accuracy. We explore the impact of speckle noise on SAR target recognition performance. Subsequently, we propose a deep attention convolutional network assisted by a multiscale residual despeckling network to improve SAR target recognition accuracy under speckle corruption. Noise information is first learned via a despeckling subnetwork consisting of dual-branch multiscale feature extraction, feature fusion, and adaptive feature channel selection. Then a classification subnetwork with a cross-dimension interaction attention mechanism is designed to realize feature extraction and identity reasoning of SAR targets. Experimental results on the benchmark moving and stationary target acquisition and recognition dataset demonstrate the effectiveness and superiority of the proposed method. (c) 2023 Society of Photo-Optical Instrumentation Engineers (SPIE)  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
1931-3195  
入藏号:  
WOS:000969337300025  
Indexed Date:  
2023-05-12  
Institution Address:  
Univ Elect Sci & Technol China, Sch Informat & Commun Engn, Chengdu, Peoples R China  
Institution:

### 357. Cross-scale feature fusion connection for a YOLO detector

摘要: Multi-scale feature fusion is often used to address the issue of scale variations in object detection. However, most of the proposed network architectures only combine the features of two adjacent levels sequentially, so the first fusion nodes in both top-down and bottom-up pathways must be blank nodes that only have one input with no feature fusion. In this work, cross-scale feature fusion connection (CFFC) is proposed which aims to enhance the entire feature hierarchy by propagating the features of each level more efficiently. The proposed method reuses and aggregates all the features of other scales to the blank nodes in both top-down and bottom-up pathways. Furthermore, the authors remove the 1 x 1 convolutional layer and replace the shortcut with concatenation before fusing multiple features. These concatenated feature maps are then supervised by the channel attention block at the fusion nodes. This modification allows the network to learn the important degree of each level in concatenated feature maps along the channel dimension. It is also observed that the proposed method alleviates the inconsistency in feature pyramids with fewer parameters. The performance of a YOLO object detector equipped with the proposed method on the COCO test-dev 2017 is evaluated. The results show that the proposed method outperforms other architectures presented in the literature.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1751-9632  
eISSN:  
1751-9640  
入藏号:  
WOS:000706173300001  
Indexed Date:  
2021-10-24  
Institution Address:  
Xian Inst Opt & Precis Mech CAS, Xian 710119, Shaanxi, Peoples R China  
Institution Address:  
China Astronaut Res & Training Ctr, Beijing, Peoples R China  
Institution:  
Institution:  
China Astronaut Res & Training Ctr

### 358. YOLO-AMM: A Real-Time Classroom Behavior Detection Algorithm Based on

摘要: Classroom behavior detection is a key task in constructing intelligent educational environments. However, the existing models are still deficient in detail feature capture capability, multi-layer feature correlation, and multi-scale target adaptability, making it challenging to realize high-precision real-time detection in complex scenes. This paper proposes an improved classroom behavior detection algorithm, YOLO-AMM, to solve these problems. Firstly, we constructed the Adaptive Efficient Feature Fusion (AEFF) module to enhance the fusion of semantic information between different features and improve the model's ability to capture detailed features. Then, we designed a Multi-dimensional Feature Flow Network (MFFN), which fuses multi-dimensional features and enhances the correlation information between features through the multi-scale feature aggregation module and contextual information diffusion mechanism. Finally, we proposed a Multi-Scale Perception and Fusion Detection Head (MSPF-Head), which significantly improves the adaptability of the head to different scale targets by introducing multi-scale feature perception, feature interaction, and fusion mechanisms. The experimental results showed that compared with the YOLOv8n model, YOLO-AMM improved the mAP0.5 and mAP0.5-0.95 by 3.1% and 4.0%, significantly improving the detection accuracy. Meanwhile, YOLO-AMM increased the detection speed (FPS) by 12.9 frames per second to 169.1 frames per second, which meets the requirement for real-time detection of classroom behavior.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001431624100001  
PubMed ID:  
40006371  
Indexed Date:  
2025-03-03  
Institution Address:  
Wuxi Univ, Sch Internet Things Engn, Wuxi 214105, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Automat, Nanjing 210044, Peoples R China  
Institution Address:  
Jiangsu Foreign Expert Lab, Wuxi 214105, Peoples R China  
Institution:  
Institution:  
Institution:  
Jiangsu Foreign Expert Lab

### 359. Expansive Receptive Field and Local Feature Extraction Network:

摘要: Fibroadenoma is a common benign breast disease that affects women of all ages. Early diagnosis can greatly improve the treatment outcomes and reduce the associated pain. Computer-aided diagnosis (CAD) has great potential to improve diagnosis accuracy and efficiency. However, its application in sonography is limited. A network that utilizes expansive receptive fields and local information learning was proposed for the accurate segmentation of breast fibroadenomas in sonography. The architecture comprises the Hierarchical Attentive Fusion module, which conducts local information learning through channel-wise and pixel-wise perspectives, and the Residual Large-Kernel module, which utilizes multiscale large kernel convolution for global information learning. Additionally, multiscale feature fusion in both modules was included to enhance the stability of our network. Finally, an energy function and a data augmentation method were incorporated to fine-tune low-level features of medical images and improve data enhancement. The performance of our model is evaluated using both our local clinical dataset and a public dataset. Mean pixel accuracy (MPA) of 93.93% and 86.06% and mean intersection over union (MIOU) of 88.16% and 73.19% were achieved on the clinical and public datasets, respectively. They are significantly improved over state-of-the-art methods such as SegFormer (89.75% and 78.45% in MPA and 83.26% and 71.85% in MIOU, respectively). The proposed feature extraction strategy, combining local pixel-wise learning with an expansive receptive field for global information perception, demonstrates excellent feature learning capabilities. Due to this powerful and unique local-global feature extraction capability, our deep network achieves superior segmentation of breast fibroadenoma in sonography, which may be valuable in early diagnosis.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2948-2925  
eISSN:  
2948-2933  
入藏号:  
WOS:001236098500001  
PubMed ID:  
38822159  
Indexed Date:  
2024-06-06  
Institution Address:  
Chongqing Med Univ, Coll Biomed Engn, State Key Lab Ultrasound Med & Engn, Med Coll Rd, Chongqing 400016, Peoples R China  
Institution Address:  
Chongqing Med Univ, Chongqing Key Lab Biomed Engn, Chongqing 400016, Peoples R China  
Institution Address:  
Natl Med Prod Adm NMPA Key Lab Qual Evaluat Ultras, Key Lab Qual Evaluat, Ultrason Surg Equipment, Donghu New Technol Dev Zone, 507 Gaoxin Ave, Wuhan 430075, Hubei, Peoples R China  
Institution:  
Institution:  
Institution:  
Natl Med Prod Adm NMPA Key Lab Qual Evaluat Ultras

### 360. A Small Target Pedestrian Detection Model Based on Autonomous Driving

摘要: Since small-target pedestrians account for a small proportion of pixels in images and lack texture features, the feature information of small-target pedestrians is often ignored in the feature extraction process, leading to reduced accuracy and poor robustness. To improve the accuracy of small-target pedestrian detection and the anti-interference ability of the model, a small-target pedestrian detection model that fuses residual networks and feature pyramids is proposed. First, a residual block with a discard layer is constructed to replace the standard residual block in the residual network structure to reduce the complexity of the model computation process and solve the problems of gradient disappearance and explosion in the deep network. Then, feature selection and feature alignment modules are added to the lateral connection part of the feature pyramid to enhance important pedestrian features in the input image, and the multiscale feature fusion capability of the model is enhanced for small-target pedestrians, thereby improving the detection accuracy of small-target pedestrians and solving the problems of feature misalignment and ignored multiscale features in the feature pyramid network. Finally, a cascaded autofocus query module is proposed to increase the inference speed of the feature pyramid network through focusing and querying, thus improving the performance and efficiency of small-target pedestrian detection. The experimental results show that the proposed model achieves better detection results than previous models.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0197-6729  
eISSN:  
2042-3195  
入藏号:  
WOS:000943327100001  
Indexed Date:  
2023-03-18  
Institution Address:  
Fujian Univ Technol, Sch Transportat, Fuzhou, Fujian, Peoples R China  
Institution Address:  
Fujian Univ Technol, Intelligent Transportat Syst Res Ctr, Fuzhou, Fujian, Peoples R China  
Institution:  
Institution:

### 361. Superresolution reconstruction of optical remote sensing images based on

摘要: Due to the influence of imaging equipment and environmental conditions on optical remote sensing image acquisition, image resolution is generally low. Superresolution reconstruction technology is an important way to improve image quality. However, the existing optical remote sensing image superresolution reconstruction methods have some problems, such as insufficient feature extraction, blurred texture details of reconstructed images, and excessive network accumulation. To solve the above problems, a superresolution reconstruction method for optical remote sensing images based on a multiscale attention adversarial network is proposed in this paper. The method takes a generative adversarial network (GAN) as the basic framework. The generator uses four multiscale attention residual blocks (MSARBs) to extract image multiscale feature information and carries out feature fusion through a binary feature fusion structure (BFFS) to generate more realistic images. The discriminator uses a depth convolution network to distinguish the differences between real images and superresolution images. In the aspect of loss function construction, the perceptual loss and adversarial loss are combined to improve the perceptual quality of the images. Experimental results show that this method is superior to the compared algorithm in regard to the objective evaluation metrics of peak signal-to-noise ratio (PSNR) and structural similarity (SSIM), and its reconstructed images have better visual effect.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000778916300006  
Indexed Date:  
2022-04-21  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao 266590, Peoples R China  
Institution Address:  
Shandong Univ Sci & Technol, Shandong Prov Key Lab Wisdom Mine Informat Techno, Qingdao 266590, Peoples R China  
Institution:  
Institution:

### 362. BMSeNet: Multiscale Context Pyramid Pooling and Spatial Detail

摘要: Most real-time semantic segmentation networks use shallow architectures to achieve fast inference speeds. This approach, however, limits a network's receptive field. Concurrently, feature information extraction is restricted to a single scale, which reduces the network's ability to generalize and maintain robustness. Furthermore, loss of image spatial details negatively impacts segmentation accuracy. To address these limitations, this paper proposes a Multiscale Context Pyramid Pooling and Spatial Detail Enhancement Network (BMSeNet). First, to address the limitation of singular semantic feature scales, a Multiscale Context Pyramid Pooling Module (MSCPPM) is introduced. By leveraging various pooling operations, this module efficiently enlarges the receptive field and better aggregates multiscale contextual information. Moreover, a Spatial Detail Enhancement Module (SDEM) is designed, to effectively compensate for lost spatial detail information and significantly enhance the perception of spatial details. Finally, a Bilateral Attention Fusion Module (BAFM) is proposed. This module leverages pixel positional correlations to guide the network in assigning appropriate weights to the features extracted from the two branches, effectively merging the feature information of both branches. Extensive experiments were conducted on the Cityscapes and CamVid datasets. Experimental results show that the proposed BMSeNet achieves a good balance between inference speed and segmentation accuracy, outperforming some state-of-the-art real-time semantic segmentation methods.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
1424-8220  
入藏号:  
WOS:001304642300001  
PubMed ID:  
39204841  
Indexed Date:  
2024-09-08  
Institution Address:  
Henan Polytech Univ, Sch Software, Jiaozuo 454000, Peoples R China  
Institution:

### 363. Semi-Supervised Soft Sensor Modeling Based on Ensemble Learning With

摘要: Key quality variables are critical in the industrial production process. The soft sensor technology, which predicts key quality variables by establishing mathematical models, has gradually become a research hotspot. However, due to the difficulty in obtaining labeled data in industrial fields, a substantial quantity of unlabeled data is not reasonably utilized, which challenges the reliability and accuracy of conventional soft sensor models. Therefore, a semi-supervised model based on the voting ensemble learning is proposed, which combines the outcomes of multiple models' predictions and utilizes a genetic optimization algorithm to iteratively optimize the generated pseudolabels, improving the accuracy of pseudolabels. By using the channel attention mechanism and multiscale feature fusion method, the feature extraction ability of the model is further improved, thus enhancing the prediction accuracy of the model. Finally, experiments were carried out on industrial debutanizer and industrial steam volume datasets to validate the superior predictive performance of the proposed method.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001276438700038  
Indexed Date:  
2024-08-08  
Institution Address:  
Northwest Normal Univ, Coll Comp Sci & Engn, Lanzhou 730070, Peoples R China  
Institution:

### 364. A temperature measurement compensation method for industrial rotary

摘要: Accurate temperature measurement is vital for assessing the reaction atmosphere of industrial rotary kilns. However, dynamic water mist interference imposes significant challenges to accurate infrared temperature measurement. To this end, this study proposes a temperature measurement compensation method based on infrared multi-feature fusion. First, we developed an industrial system to collect infrared data from the kiln. To address the issue of unclear water mist in infrared images caused by its low resolution and weak texture features, we designed an artificial feature based on infrared images and temperature differences named water mist level (WML). This feature effectively represents the size of the water mist in the image. Additionally, a classification model is established to identify the WML automatically. Subsequently, a novel multi-scale feature fusion network called efficient regression feature pyramid network (ERFPN) is proposed to acquire multi-scale image features. Finally, we propose a grouped feature fusion network (GFFN) to fuse multi-scale image features, WML, and interfered temperature. Experimental results show that the proposed compensation method achieves satisfactory infrared temperature measurement results and significantly reduces the temperature measurement error caused by water mist.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1350-4495  
eISSN:  
1879-0275  
入藏号:  
WOS:001298114000001  
Indexed Date:  
2024-08-30  
Institution Address:  
Cent South Univ, Sch Automat, Changsha 410083, Peoples R China  
Institution:

### 365. Optimized anchor-free network for dense rotating object detection in

摘要: Extracting dense rotating objects accurately from remote sensing images is an emerging task in object detection. To increase the applicability of existing algorithms in the above tasks, an optimized anchor-free network optimized by a dual attention mechanism (DAM) and gate multiscale feature fusion (GMFF) is designed. The DAM module is composed of two attention mechanisms with different functions. This part can enhance the backbone network's ability to extract and model information at different levels and reduce the accuracy loss caused by object density changes in the image. The GMFF module uses the gating structure to realize adaptive transmission and integration of multiscale information. Through this module, the useless information in features will be filtered, and the key information will be retained. Several experiments are designed to verify the feasibility of the algorithm. Compared with the baseline model, adding DAM and GMFF to the dense rotating object extraction task in remote sensing images improves the model accuracy by 3.5% and 2.1%, respectively, while adding two modules simultaneously, and the accuracy increases from 79.1% to 84.3%. In conventional object extraction tasks, such as dataset for object detection in aerial images and HRSC2016, our method has the highest accuracy compared to other similar algorithms, with 76.5% and 90.3%, respectively.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:001172000900013  
Indexed Date:  
2024-03-17  
Institution Address:  
Key Lab Ecol Environm Protect Space Informat Appli, Zhengzhou, Peoples R China  
Institution Address:  
China Elect Technol Grp Corp, Res Inst 27, Zhengzhou, Peoples R China  
Institution:  
Key Lab Ecol Environm Protect Space Informat Appli  
Institution:

### 366. Research on Rapid Recognition of Moving Small Targets by Robotic Arms

摘要: For small target objects on fast-moving conveyor belts, traditional vision detection algorithms equipped with conventional robotic arms struggle to capture the long and short-range pixel dependencies crucial for accurate detection. This leads to high miss rates and low precision. In this study, we integrate the traditional EMA (efficient multi-scale attention) algorithm with the c2f (channel-to-pixel) module from the original YOLOv8, alongside a Faster-Net module designed based on partial convolution concepts. This fusion results in the Faster-EMA-Net module, which greatly enhances the ability of the algorithm and robotic technologies to extract pixel dependencies for small targets, and improves perception of dynamic small target objects. Furthermore, by incorporating a small target semantic information enhancement layer into the multiscale feature fusion network, we aim to extract more expressive features for small targets, thereby boosting detection accuracy. We also address issues with training time and subpar performance on small targets in the original YOLOv8 algorithm by improving the loss function. Through experiments, we demonstrate that our attention-based visual detection algorithm effectively enhances accuracy and recall rates for fast-moving small targets, meeting the demands of real industrial scenarios. Our approach to target detection using industrial robotic arms is both practical and cutting-edge.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
2  
eISSN:  
2076-3417  
入藏号:  
WOS:001233120000001  
Indexed Date:  
2024-06-03  
Institution Address:  
Southwest Univ, Sch Engn & Technol, Chongqing 400715, Peoples R China  
Institution:

### 367. VKP-P3D: Real-Time Monocular Pseudo 3D Object Detection Based on Visible

摘要: Three-dimensional object detection has been substantially improved with the use of expensive LiDAR and stereo vision systems in intelligent driving. The less-expensive and more scalable solution of monocular 3D object detection, however, remains a key challenge. This study primarily explores real-time pseudo 3D object detection with monocular vision and designs a single-shot RPN model, VKP-P3D, which relies purely on visual feature extraction. Through a multiscale feature fusion and an attention mechanism module, this model obtains high-dimensional feature representations during the feature extraction phase. In the detection head of the VKP-P3D model, the pseudo 3D object detection is obtained by regressing 2D bounding box and the visible key points within the image coordinate of the 3D box from the camera's perspective. Finally, assuming flat ground and considering geometric parameters of the camera, the object's 3D information can be extracted. To verify the effectiveness of the proposed algorithm, we constructed two pseudo 3D object detection datasets based on visible key points and compared with current state-of-the-art real-time object detector. Results showed that the proposed model has high detection accuracy and speed.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001192240200001  
Indexed Date:  
2024-04-03  
Institution Address:  
Hunan Univ, Coll Elect & Informat Engn, Changsha 410082, Peoples R China  
Institution:

### 368. Edge-Guided Feature Pyramid Networks: An Edge-Guided Model for Enhanced

摘要: Infrared small target detection technology has been widely applied in the defense sector, including applications such as precision targeting, alert systems, and naval monitoring. However, due to the small size of their targets and the extended imaging distance, accurately detecting drone targets in complex infrared environments remains a considerable challenge. Detecting drone targets accurately in complex infrared environments poses a substantial challenge. This paper introduces a novel model that integrates edge characteristics with multi-scale feature fusion, named Edge-Guided Feature Pyramid Networks (EG-FPNs). This model aims to capture deep image features while simultaneously emphasizing edge characteristics. The goal is to resolve the problem of missing target information that occurs when Feature Pyramid Networks (FPNs) perform continuous down-sampling to obtain deeper semantic features. Firstly, an improved residual block structure is proposed, integrating multi-scale convolutional feature extraction and inter-channel attention mechanisms, with significant features being emphasized through channel recalibration. Then, a layered feature fusion module is introduced to strengthen the shallow details in images while fusing multi-scale image features, thereby strengthening the shallow edge features. Finally, an edge self-fusion module is proposed to enhance the model's depiction of image features by extracting edge information and integrating it with multi-scale features. We conducted comparative experiments on multiple datasets using the proposed algorithm and existing advanced methods. The results show improvements in the IoU, nIoU, and F1 metrics, while also showcasing the lightweight nature of EG-FPNs, confirming that they are more suitable for drone detection in resource-constrained infrared scenarios.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001377731100001  
PubMed ID:  
39686304  
Indexed Date:  
2024-12-21  
Institution Address:  
Nanjing Univ Sci & Technol, Natl Key Lab Transient Phys, Nanjing 210094, Peoples R China  
Institution Address:  
Nanjing Univ Sci & Technol, Sch Elect & Opt Engn, Nanjing 210094, Peoples R China  
Institution:  
Institution:

### 369. Multi-level feature enhancement network for object detection in sonar

摘要: The unstable geometric features affect the accuracy of object detection in sonar images. We thus propose a novel multi-level feature enhancement network to enhance useful features for object detection in sonar images. We first introduce a deformable convolution to model variations in geometric features. In addition, spatial and channel attention modules are utilized to aggregate rich semantic information from features, improving the quality of feature extraction. We further use an adaptive multi-scale feature fusion module for feature weighting so as to enhance fine-grained features and minimize information loss during feature fusion. Then, the cascaded detection module corrects the prediction results of the previous detector with a low Intersection-over-Union (IoU) threshold, where each detector employs adaptive feature enhancement blocks to enhance region proposal features and thus improve detection performance. Experimental results on two real-world sonar image datasets show that our proposed model performs better than several mainstream object detection methods by achieving 2% to 19.4% higher accuracy rates.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1047-3203  
eISSN:  
1095-9076  
入藏号:  
WOS:001239642600001  
Indexed Date:  
2024-06-10  
Institution Address:  
Dalian Maritime Univ, Sch Informat Sci & Technol, Dalian 116026, Peoples R China  
Institution Address:  
East China Normal Univ, Sch Data Sci & Engn, Shanghai 200062, Peoples R China  
Institution Address:  
Dalian Maritime Univ, Coll Marine Engn, Dalian 116026, Peoples R China  
Institution:  
Institution:  
Institution:

### 370. I-YOLO: a novel single-stage framework for small object detection

摘要: Small object detection is a challenging task in computer vision. We claim that the huge performance gap between the small object detectors and normal sized object detectors stems from two aspects, including the small object dataset and the small object itself. In terms of datasets, we build a large-scale dataset with high image resolution dubbed Small-PCB, in order to promote detection in semiconductor industry. For the small object itself, we utilize multi-scale feature learning and feature fusion strategy to help detect objects. More concretely, we devise two novel components to predict small objects better: re-parameterized module with channel shuffle (RMCS) and multi-scale feature enhanced convolution (MFEC). MFEC aims to split input channels into several parts and applies convolutions with different sizes to each part, and adopt point-by-point convolution to fuse individual channel features. RMCS not only use structural re-parameterization, but also channel shuffle. The usage of channel shuffle can be seen as a fusion of channel features. It strengthens feature information interaction between different channel groups, which bring more informative feature clues. Based on the RMCS and the MFEC, we introduce OIU-RMCS and M-MFEC, respectively. Finally, we build our I-YOLO via integrating these two components into a YOLO-based detector. A large number of qualitative and quantitative results in the experiments indicate that our proposed I-YOLO achieves the state-of-the-art performance on the popular AI-TODv2 and Small-PCB datasets.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
4  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001168975900001  
Indexed Date:  
2024-03-04  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Elect & Informat Engn, Nanjing 211106, Peoples R China  
Institution:

### 371. Model With Leader-Follower Backbone and Bifurcation Fusion for UAV

摘要: The application of unmanned aerial vehicles (UAVs) is crucial in traffic information collection. In addressing the challenge of detecting small targets in UAV imagery, simply increasing the model depth is not the most optimal solution. In this work, we propose a leader-follower network with bifurcation fusion for detection (MSDet), a novel object detection method based on leader-follower backbone and bifurcation fusion. Different feature extraction methods provide varying feature information, and their fusion enables a more comprehensive and multidimensional description of the target. The simplified auxiliary networks are connected layer by layer with the main backbone, and their final output is fed back to the initial feature map. The main backbone and auxiliary networks can be flexibly selected and combined to adapt to the unique features of different scenarios. Bifurcation fusion achieves flexible multiscale feature fusion by introducing branches during the top-down fusion process. One branch performs deeper top-down fusion to capture more shallow features, whereas the opposing branch offers a comprehensive understanding of the overall structure. Experimental results demonstrate that the proposed method outperforms state-of-the-art methods when applied to three UAV datasets. Furthermore, this study suggests that integrating with different backbones may yield better performance than simply scaling up models when faced with challenging situations.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001410586700011  
Indexed Date:  
2025-02-13  
Institution Address:  
Shandong Jiaotong Univ, Sch Rail Transportat, Jinan 250353, Shandong, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Multimodal Artificial Intelligence S, Beijing 100045, Peoples R China  
Institution Address:  
Chinese Acad Sci, Guangdong Engn Res Ctr 3 D Printing & Intelligent, Cloud Comp Ctr, Dongguan 523803, Peoples R China  
Institution:  
Institution:  
Institution:

### 372. Joint Detection and Association for End-to-End Multi-object Tracking

摘要: Multi-object tracking (MOT) is mainly used for detecting and tracking the object on multi-cameras, which is widely applied in intelligent video surveillance and intelligent security. The process of MOT generally involves three import parts: feature extracting, multi-task learning and object matching. Unfortunately, the existed methods still have some drawbacks. Firstly, the feature extracting module cannot effectively fuse the shallow and deep features. What's more, the multi-task learning module cannot strike a good balance between the detection and re-identification. In addition, the object matching module associates with pedestrian by using a traditional method rather than training a model. For these problems, we propose a method of joint detection and association (JDA) for end-to-end multi-object tracking network, which involves the multi-scale feature extraction and the learnable object association. It first combines a feature extraction backbone based on multi-scale feature fusion and a point-based multi-task object detection branch, to solve the task of feature extraction and object detection. Then, a learnable object motion association module is embedded, which uses the historical frames information to infer the position of the object, and associate the object identity between previous frames and subsequent frames. In addition, the JDA can be end-to-end trained when handling the detection and matching tasks. The proposed JDA is evaluated through a series of experiments on MOT16 and MOT17. The results shows that JDA the existing methods in terms of precision and stability of MOT.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1370-4621  
eISSN:  
1573-773X  
入藏号:  
WOS:001118759500001  
Indexed Date:  
2023-12-21  
Institution Address:  
Shenzhen Inst Informat Technol, Shenzhen, Peoples R China  
Institution Address:  
Univ Elect Sci & Technol China, Chengdu, Peoples R China  
Institution Address:  
Kashi Inst Elect & Informat Ind, Varanasi, Peoples R China  
Institution Address:  
Shenzhen Univ, Shenzhen, Peoples R China  
Institution:  
Institution:  
Institution:  
Kashi Inst Elect & Informat Ind  
Institution:

### 373. Few-Shot PCB Surface Defect Detection Based on Feature Enhancement and

摘要: In printed circuit board (PCB) defect detection, it is difficult to collect defect samples, and the detection effect is poor due to the lack of data. On the basis of the few-shot learning method, a few-shot PCB defect detection model is proposed. This model introduces feature enhancement module and multi-scale fusion module. The feature enhancement module based on the improved convolution block attention module (CBAM) can highlight the key areas of the received feature maps and suppress the interference of useless information. Aiming at the small size of PCB defects, a multi-scale feature fusion strategy is proposed. It can extract multi-scale feature maps of PCB and fuse them into a high-quality feature map containing different scale information, which can improve the detection precision of the model for small object defects. A large number of experiments on PCB dataset show that our few-shot PCB defect detection model outperforms state-of-the-art methods under different shot settings (k=1,2,3,5,10,30). Notably, the proposed model can take into account both detection efficiency and precision, which means it has high practical application value.  
Web of Science 核心合集中的 "被引频次":  
13  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
13  
ISSN:  
2169-3536  
入藏号:  
WOS:000902046100001  
Indexed Date:  
2023-01-06  
Institution Address:  
Taiyuan Univ Technol, Coll Elect & Power Engn, Taiyuan 030024, Peoples R China  
Institution Address:  
Taiyuan Univ Technol, Coll Informat & Comp, Jinzhong 030600, Peoples R China  
Institution:  
Institution:

### 374. Retinal vessel segmentation based on multi-scale attention feature

摘要: To solve the problem of irregular and difficult segmentation of blood vessels in fundus retinal images, a multi-scale attention feature fusion network model based on a dual-decoder structure is proposed to achieve accurate segmentation of retinal blood vessels. The dual decoder branch network structure can reduce information loss. In the encoder, the multi-scale attention feature fusion module is designed to extract rich multi-scale features and the spatial attention module is combined to enhance the extraction of spatial context information and improve vascular recognition ability. Squeeze-and-excitation module is used to optimize aggregated features, suppress irrelevant feature channels and improve the comprehensive segmentation ability of the model. The experimental results on the DRIVE and CHASEDB1 data sets show that the recall rate can reach 0.841 1 and 0.855 1 respectively, making great progress compared with some advanced networks at present, with the maximum increase of 6.6% and 8.25% respectively.  
摘要:  
针对眼底视网膜图像中血管形态不规则、难以分割的问题,提出一种基于双解码器结构的多尺度注意力特征融合网络模型,可以实现视网膜血管精确分割。双解码器分支网络结构能减少信息丢失,编码器中设计多尺度注意力特征融合模块来提取丰富的多尺度特征,结合空间注意力模块加强空间上下文信息提取,提高血管识别能力。利用挤压与激励模块对融合特征进行优化,抑制不相关特征通道,提高模型综合分割能力。在DRIVE和CHASEDB1数据集上的实验结果显示,召回率分别达到0.841 1和0.855 1,相较目前一些先进网络取得了较大进步,最大提升分别达到6.6%和8.25%。  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1007-130X  
入藏号:  
CSCD:7618799  
Indexed Date:  
2024-03-02  
Institution Address:  
湖南师范大学信息科学与工程学院, 长沙, 湖南 410081, 中国  
Institution Address:  
College of Information Science and Engineering,Hunan Normal University, Changsha, Hunan 410081, China  
Institution:  
湖南师范大学信息科学与工程学院  
Institution:  
College of Information Science and Engineering,Hunan Normal University

### 375. A Multi-scale Fusion Obstacle Detection Algorithm for Autonomous Driving

摘要: Effective circumstance perception technology is the prerequisite for the successful application of autonomous driving, especially the detection technology of traffic objects that affects other tasks such as driving decisions and motion execution in autonomous vehicles. However, recent studies show that a single sensor cannot perceive the surrounding environment stably and effectively in complex circumstances. In the article, we propose a multi-scale feature fusion framework that exploits a dual backbone network to extract camera and radar feature maps and performs feature fusion on three different feature scales using a new fusion module. In addition, we introduce a new generation mechanism of radar projection images and relabel the nuScenes dataset since there is no other suitable autonomous driving dataset for model training and testing. The experimental results show that the fusion models achieve superior accuracy over visual image-based models on the evaluation criteria of PASCAL visual object classes (VOC) and Common Objects in Context (COCO), about 2% over the baseline model (YOLOX).  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2574-0741  
eISSN:  
2574-075X  
入藏号:  
WOS:001274850200005  
Indexed Date:  
2024-08-01  
Institution Address:  
Hunan Univ, State Key Lab Adv Design & Mfg Vehicle Body, Changsha, Peoples R China  
Institution:

### 376. A Channel-Wise Multi-Scale Network for Single Image Super-Resolution

摘要: Existing multi-scale feature extraction methods extract image features using various convolution window sizes conducted on the spatial dimension of the feature maps. However, such an approach inevitably encounters redundant convolution operations. To address this concern, we propose to extract multi-scale features on the channel dimension rather than on the spatial dimension. To demonstrate, a channel-wise multi-scale network (CMSN) is proposed for conducting single image super-resolution (SISR). In our CMSN, a sequence of channel-wise multi-scale blocks (CMSBs) is designed to extract multi-scale features at increasing levels by performing convolutions with different channel numbers (i.e., scales). To fuse the image features generated from different levels in our CMSN, a hybrid attention-aware feature fusion block (HAFFB) is proposed. Extensive experimental results have clearly shown the superiority of our CMSN to that of several state-of-the-art SISR methods on delivering superior high-resolution images, both objectively and subjectively. This reveals the potential of channel-wise, versus spatial-wise, on the effectiveness of multi-scale feature extraction.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1070-9908  
eISSN:  
1558-2361  
入藏号:  
WOS:001188292700013  
Indexed Date:  
2024-04-07  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Elect & Informat Engn, Nanjing 210023, Peoples R China  
Institution Address:  
Soochow Univ, Sch Comp Sci & Technol, Suzhou 215006, Peoples R China  
Institution:  
Institution:

### 377. EC-PFN: a multiscale woven fusion network for industrial product surface

摘要: In order to address challenges such as small target sizes, low contrast, significant intra-class variations, and indistinct inter-class differences in surface defect detection, this paper proposes the Enhanced Context-aware Parallel Fusion Network (EC-PFN). The network employs a Featur Weave Network architecture to enhance contextal awareess and parallel fusion capabilities. It utilizes a Feature Fusion Module (UniFusionLayer) for effective multiscale and multisemantic feature learning, offering new perspectives on feature fusion. Additionally, a Receptive Field Block (RFB) module is introduced to expand the receptive field, enhancing feature extraction in scenarios with low contrast and subtle defects. The Loss Ranking Module (LRM) is incorporated to optimize the target-oriented loss, improving performance by omitting low-confidence bounding boxes. Extensive experiments on a light guide plate defect dataset demonstrate that EC-PFN achieves a detection accuracy (mAP) of 98.9%, a detection speed of 92 FPS, and a computational cost of 14.5 GFLOPs, outperforming mainstream surface defect detection models.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2199-4536  
eISSN:  
2198-6053  
入藏号:  
WOS:001375690400002  
Indexed Date:  
2024-12-18  
Institution Address:  
Zhejiang Sci Tech Univ, Sch Mat Sci & Engn, Hangzhou 310018, Zhejiang, Peoples R China  
Institution:

### 378. Feature Enhancement Based on CycleGAN for Nighttime Vehicle Detection

摘要: Existing night vehicle detection methods mainly detect vehicles by detecting headlights or taillights. However, these features are adversely affected by the complex road lighting environment. In this paper, a cascade detection network framework FteGanOd is proposed with a feature translate-enhancement (FTE) module and the object detection (OD) module. First, the FTE module is built based on CycleGAN and multi-scale feature fusion is proposed to enhance the detection of vehicle features at night. The features of night and day are combined by fusing different convolutional layers to produce enhanced feature (EF) maps. Second, the OD module, based on the existing object detection network, is improved by cascading with the FTE module to detect vehicles on the EF maps. The proposed FteGanOd method recognizes vehicles at night with greater accuracy by improving the contrast between vehicles and the background and by suppressing interference from ambient light. The proposed FteGanOd is validated on the Berkeley Deep Drive (BDD) dataset and our private dataset. The experimental results show that our proposed method can effectively enhance vehicle features and improve the accuracy of vehicle detection at night.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
30  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
30  
ISSN:  
2169-3536  
入藏号:  
WOS:000606549600001  
Indexed Date:  
2021-07-21  
Institution Address:  
Beijing Jiaotong Univ, Sch Elect & Informat Engn, Beijing 100044, Peoples R China  
Institution:

### 379. Improved real-time infrared small target detection based on YOLOv5s

摘要: In this paper,an improved infrared small target detection model,infrared-YOLOv5s,based on YOLOv5s is proposed to address the problems of low resolution,complex background and lack of detailed features of infrared images.In feature extraction stage,SPD-Conv is used for down-sampling,which divides the feature map into feature submaps and concatenate them by channel to avoid the loss of features caused by down-sampling in the process of multiscale feature extraction.And an improved atrous spatial pyramid pooling module is designed to improve feature extraction capabilities by fusing features with different receptive fields.Then,in feature fusion stage,a deep-to-shallow attention module is introduced to embed deep semantic features into shallow spatial features to enhance the expression of shallow features.Moreover,in prediction stage,the prediction layers,feature extraction layers and feature fusion layers for large target detection in the network are cut down to reduce the model size and improve real-time performance at the same time.The effectiveness of each module is verified by ablation experiments,and experimental results show that the proposed model achieves 95.4 % mAP0.5 of on SIRST dataset,which is 2.3 % higher than that of original YOLOv5s.The model size is reduced by 72.9 % to 4.5 MB,and the inference speed on Nvidia Xavier reaches 28 f/s,which is conducive to the actual deployment and application.Therefore,the effectiveness of the proposed model is further verified by transfer experiments using Infrared-PV dataset,and the proposed model can meet the real-time requirements while improving the performance of small target detection in infrared images,and is suitable for the task of real-time small target detection in infrared images.  
摘要:  
针对红外图像分辨率低、背景复杂、目标细节特征缺失等问题,提出了一种基于YOLOv5s的改进实时红外小目标检测模型Infrared-YOLOv5s。在特征提取阶段,采用SPDConv进行下采样,将特征图切分为特征子图并按通道拼接,避免了多尺度特征提取过程中下采样导致的特征丢失情况,设计了一种基于空洞卷积的改进空间金字塔池化模块,通过对具有不同感受野的特征进行融合来提高特征提取能力;在特征融合阶段,引入由深到浅的注意力模块,将深层特征语义特征嵌入到浅层空间特征中,增强浅层特征的表达能力;在预测阶段,裁减了网络中针对大目标检测的特征提取层、融合层及预测层,降低模型大小的同时提高了实时性。首先通过消融实验验证了提出各模块的有效性,实验结果表明,改进模型在SIRST数据集上平均精度均值达到了95.4 %,较原始YOLOv5s提高了2.3 %,且模型大小降低了72.9 %,仅为4.5 M,在Nvidia Xavier上推理速度达到28 f/s,利于实际的部署和应用。在Infrared-PV数据集上的迁移实验进一步验证了改进算法的有效性。提出的改进模型在提高红外图像小目标检测性能的同时,能够满足实时性要求,因而适用于红外图像小目标实时检测任务。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1001-5078  
入藏号:  
CSCD:7698756  
Indexed Date:  
2024-06-29  
Institution Address:  
杭州电子科技大学自动化学院, 杭州, 浙江 310018, 中国  
Institution Address:  
杭州电子科技大学圣光机联合学院, 杭州, 浙江 310018, 中国  
Institution Address:  
School of Automation,Hangzhou Dianzi University, Hangzhou, Zhejiang 310018, China  
Institution Address:  
HDU-ITMO Joint Institute,Hangzhou Dianzi University, Hangzhou, Zhejiang 310018, China  
Institution:  
杭州电子科技大学自动化学院  
Institution:  
杭州电子科技大学圣光机联合学院  
Institution:  
School of Automation,Hangzhou Dianzi University  
Institution:  
HDU-ITMO Joint Institute,Hangzhou Dianzi University

### 380. Crowd counting network based on multi-scale spatial attention feature

摘要: Concerning the poor performance problem of crowd counting tasks in different dense scenes caused by severe scale changes and occlusions,a new Multi-scale spatial Attention Feature fusion Network(MAFNet)was proposed based on the Congested Scene Recognition Network(CSRNet)by combining the multi-scale feature fusion structure and the spatial attention module.Before extracting features with MAFNet,the scene images with head markers were processed with the Gaussian filter to obtain the ground truth density maps of images.In addition,the method of jointly using two basic loss functions was proposed to constrain the consistency of the density estimation map and the ground truth density map.Next,with the multi-scale feature fusion structure as the backbone of MAFNet,the strategy of extracting and fusing multi-scale features simultaneously was used to obtain the multi-scale fusion feature map,then the feature maps were calibrated and refused by the spatial attention module.After that,an estimated density image was generated through dilated convolution,and the number of people in the scene was obtained by integrating the estimated density image pixel by pixel.To verify the effectiveness of the proposed model,evaluations were conducted on four datasets (ShanghaiTech,UCF\_CC\_50,UCF\_QRNF and World-Expo'10).Experimental results on ShanghaiTech dataset PartB show that,compared with CSRNet,MAFNet has a Mean Absolute Error(MAE)reduction of 34.9% and a Mean Square Error(MSE)reduction of 29.4%.Furthermore,experimental results on multiple datasets show that by using the attention mechanism and multi-scale feature fusion strategy,MAFNet can extract more detailed information and reduce the impact of scale changes and occlusions.  
摘要:  
针对严重的尺度变化和遮挡导致在不同密集场景人群计数任务中性能差的问题,在密集场景识别网络(CSRNet)的基础上通过增加多尺度特征融合结构并引入空间注意力机制,提出了一种多尺度空间注意力特征融合网络(MAFNet)。在MAFNet进行特征提取之前,需要对添加了人头标记的场景图进行高斯滤波生成真实密度图;此外,MAFNet还通过联合使用两种基本损失函数的方法来约束密度估计图与真实密度图的一致性。接着,MAFNet以多尺度特征融合结构为主干,首先采用边提取多尺度特征边融合的策略得到多尺度融合特征图,然后使用空间注意力模块对特征图进行校准和再融合,之后通过扩张卷积生成密度估计图,最后对密度估计图逐像素积分得到场景中的人数。为了验证所提出模型的有效性,在四个人群计数数据集(ShanghaiTech、UCF\_CC\_50、UCF\_QRNF和World-Expo'10)上进行了评估。其中ShanghaiTech数据集PartB的实验结果显示,MAFNet与CSRNet相比,平均绝对误差(MAE)降低了34.9%,均方误差(MSE)降低了29.4%。在多个数据集上的实验结果表明,采用注意力机制和多尺度特征融合策略使 MAFNet可以提取更多细节信息,减少尺度变化和遮挡带来的影响。  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1001-9081  
入藏号:  
CSCD:6911174  
Indexed Date:  
2021-05-21  
Institution Address:  
四川大学电子信息学院, 成都, 四川 610065, 中国  
Institution Address:  
College of Electronics and Information Engineering,Sichuan University, Chengdu, Sichuan 610065, China  
Institution:  
四川大学电子信息学院  
Institution:  
College of Electronics and Information Engineering,Sichuan University

### 381. Space-time video super-resolution via multi-scale feature interpolation

摘要: The goal of Space-Time Video Super-Resolution (STVSR) is to simultaneously increase the spatial resolution and frame rate of low-resolution, low-frame-rate video. In response to the problem that the STVSR method does not fully consider the spatio-temporal correlation between successive video frames, which makes the video frame reconstruction results unsatisfactory, and the problem that the inference speed of large models is slow. This paper proposes a STVSR method based on Multi-Scale Feature Interpolation and Temporal Feature Fusion (MSITF). First, feature interpolation is performed in the low-resolution feature space to obtain the features corresponding to the missing frames. The feature is then enhanced using deformable convolution with the aim of obtaining a more accurate feature of the missing frames. Finally, the temporal alignment and global context learning of sequence frame features are performed by a temporal feature fusion module to fully extract and utilize the useful spatio-temporal information in adjacent frames, resulting in better quality of the reconstructed video frames. Extensive experiments on the benchmark datasets Vid4 and Vimeo-90k show that the proposed method achieves better qualitative and quantitative performance, with PSNR and SSIM on the Vid4 dataset improving by 0.8% and 1.9%, respectively, over the state-of-the-art two-stage method AdaCof+TTVSR, and MSITF improved by 1.2% and 2.5%, respectively, compared to single-stage method RSTT. The number of parameters decreased by 80.4% and 8.2% compared to the AdaCof+TTVSR and RSTT, respectively.We release our code at https://github.com/carpenterChina/MSITF.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001287963300002  
Indexed Date:  
2024-08-15  
Institution Address:  
Guizhou Univ, Coll Comp Sci & Technol, State Key Lab Publ Big Data, Guiyang 550025, Guizhou, Peoples R China  
Institution:

### 382. Multi-Supervised Feature Fusion Attention Network for Clouds and Shadows

摘要: Cloud and cloud shadow detection are essential in remote sensing imagery applications. Few semantic segmentation models were designed specifically for clouds and their shadows. Based on the visual and distribution characteristics of clouds and their shadows in remote sensing imagery, this paper provides a multi-supervised feature fusion attention network. We design a multi-scale feature fusion block (FFB) for the problems caused by the complex distribution and irregular boundaries of clouds and shadows. The block consists of a fusion convolution block (FCB), a channel attention block (CAB), and a spatial attention block (SPA). By multi-scale convolution, FCB reduces excessive semantic differences between shallow and deep feature maps. CAB focuses on global and local features through multi-scale channel attention. Meanwhile, it fuses deep and shallow feature maps with non-linear weighting to optimize fusion performance. SPA focuses on task-relevant areas through spatial attention. With the three blocks above, FCB alleviates the difficulties of fusing multi-scale features. Additionally, it makes the network resistant to background interference while optimizing boundary detection. Our proposed model designs a class feature attention block (CFAB) to increase the robustness of cloud detection. The network achieves good performance on the self-made cloud and shadow dataset. This dataset is taken from Google Earth and contains remote sensing imagery from several satellites. The proposed model achieved a mean intersection over union (MIoU) of 94.10% on our dataset, which is 0.44% higher than the other models. Moreover, it shows high generalization capability due to its superior prediction results on HRC\_WHU and SPARCS datasets.  
Web of Science 核心合集中的 "被引频次":  
23  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
23  
eISSN:  
2220-9964  
入藏号:  
WOS:001015196300001  
Indexed Date:  
2023-07-07  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Jiangsu Key Lab Big Data Anal Technol, CICAEET, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Forestry Univ, Coll Informat Sci & Technol, Nanjing 210037, Peoples R China  
Institution:  
Institution:

### 383. A Lightweight Detection Algorithm for Unmanned Surface Vehicles Based on

摘要: Lightweight detection methods are frequently utilized for unmanned system sensing; however, when put in complicated water surface environments, they suffer from insufficient feature fusion and decreased accuracy. This paper proposes a lightweight surface target detection algorithm with multi-scale feature fusion augmentation in an effort to improve the poor detection accuracy of lightweight detection algorithms in the mission environment of unmanned surface vehicles (USVs). Based on the popular one-stage lightweight YOLOv7-Tiny target detection algorithms, a lightweight extraction module is designed first by introducing the multi-scale residual module to reduce the number of parameters and computational complexity while improving accuracy. The Mish and SiLU activation functions are used to enhance network feature extraction. Second, the path aggregation network employs coordinate convolution to strengthen spatial information perception. Finally, the dynamic head, which is based on the attention mechanism, improves the representation ability of object detection heads without any computational overhead. According to the experimental findings, the proposed model has 22.1% fewer parameters than the original model, 15% fewer GFLOPs, a 6.2% improvement in mAP@0.5, a 4.3% rise in mAP@0.5:0.95, and satisfies the real-time criteria. According to the research, the suggested lightweight water surface detection approach includes a lighter model, a simpler computational architecture, more accuracy, and a wide range of generalizability. It performs better in a variety of difficult water surface circumstances.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
eISSN:  
2077-1312  
入藏号:  
WOS:001035892700001  
Indexed Date:  
2023-08-10  
Institution Address:  
Harbin Engn Univ, Sci & Technol Underwater Vehicle Lab, Harbin 150001, Peoples R China  
Institution:

### 384. Mamba-UAV-SegNet: A Multi-Scale Adaptive Feature Fusion Network for

摘要: Accurate semantic segmentation of high-resolution images captured by unmanned aerial vehicles (UAVs) is crucial for applications in environmental monitoring, urban planning, and precision agriculture. However, challenges such as class imbalance, small-object detection, and intricate boundary details complicate the analysis of UAV imagery. To address these issues, we propose Mamba-UAV-SegNet, a novel real-time semantic segmentation network specifically designed for UAV images. The network integrates a Multi-Head Mamba Block (MH-Mamba Block) for enhanced multi-scale feature representation, an Adaptive Boundary Enhancement Fusion Module (ABEFM) for improved boundary-aware feature fusion, and an edge-detail auxiliary training branch to capture fine-grained details. The practical utility of our method is demonstrated through its application to farmland segmentation. Extensive experiments on the UAV-City, VDD, and UAVid datasets show that our model outperforms state-of-the-art methods, achieving mean Intersection over Union (mIoU) scores of 71.2%, 77.5%, and 69.3%, respectively. Ablation studies confirm the effectiveness of each component and their combined contributions to overall performance. The proposed method balances segmentation accuracy and computational efficiency, maintaining real-time inference speeds suitable for practical UAV applications.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2504-446X  
入藏号:  
WOS:001364408300001  
Indexed Date:  
2024-12-03  
Institution Address:  
Civil Aviat Flight Univ China, Coll Air Traff Management, Guanghan 618307, Peoples R China  
Institution:

### 385. Occluded Pedestrian Re-identification Method Based on Multi-scale

摘要: Pedestrian re-identification is a crucial research task in the fields of computer vision and video surveillance. The primary challenges include occlusion, illumination variation, and complex backgrounds, which significantly undermine the robustness and generalization capabilities of existing methods. To address these issues, a pedestrian re-identification method that effectively integrates multi-scale features and enhances the attention mechanism is required. This paper proposes an occluded pedestrian re-identification method based on multi-scale feature fusion. The method introduces an AAC (Add Noise and Concatenate) module, which injects noise into the central region of the input image to enhance the model's robustness and improve its generalization ability. The model employs EfficientNetB0 and DaViT\_small as backbone networks. EfficientNetB0 processes the original input image, while DaViT\_small handles the concatenated image with noise, incorporating a transposed convolution module for upsampling the feature maps to further extract high-level features and enhance spatial resolution. Additionally, a feature blocking and global fusion module is proposed, which splits the feature maps into multiple resolutions and uses different convolutional layers to further extract and fuse global features, ultimately generating a global feature vector. This design ensures the model can extract a rich variety of features from the images, thereby significantly improving the accuracy and reliability of the pedestrian re-identification task. Experimental results demonstrate that the proposed Occluded Pedestrian Re-identification method based on Multi-Scale Feature Fusion (OPR-MSFF) outperforms RNFPR (Relation Network for Person Re-Identification) with improvements of 0.7%, 0.8%, and 1.3% in rank-1 accuracy and increases of 0.9%, 1.2%, and 1.7% in mAP on the Market1501, DukeMTMC-reID, and Occluded-DukeMTMC datasets, respectively. These results validate the effectiveness of the proposed method in enhancing pedestrian re-identification performance.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1816-093X  
eISSN:  
1816-0948  
入藏号:  
WOS:001374721100018  
Indexed Date:  
2024-12-18  
Institution Address:  
Univ Sci & Technol Liaoning, Sch Comp Sci & Software Engn, Anshan 114051, Peoples R China  
Institution:

### 386. A Lightweight CER-YOLOv5s Algorithm for Detection of Construction

摘要: In the context of power-line scenarios characterized by complex backgrounds and diverse scales and shapes of targets, and addressing issues such as large model parameter sizes, insufficient feature extraction, and the susceptibility to missing small targets in engineering-vehicle detection tasks, a lightweight detection algorithm termed CER-YOLOv5s is firstly proposed. The C3 module was restructured by embedding a lightweight Ghost bottleneck structure and convolutional attention module, enhancing the model's ability to extract key features while reducing computational costs. Secondly, an E-BiFPN feature pyramid network is proposed, utilizing channel attention mechanisms to effectively suppress background noise and enhance the model's focus on important regions. Bidirectional connections were introduced to optimize the feature fusion paths, improving the efficiency of multi-scale feature fusion. At the same time, in the feature fusion part, an ERM (enhanced receptive module) was added to expand the receptive field of shallow feature maps through multiple convolution repetitions, enhancing the global information perception capability in relation to small targets. Lastly, a Soft-DIoU-NMS suppression algorithm is proposed to improve the candidate box selection mechanism, addressing the issue of suboptimal detection of occluded targets. The experimental results indicated that compared with the baseline YOLOv5s algorithm, the improved algorithm reduced parameters and computations by 27.8% and 31.9%, respectively. The mean average precision (mAP) increased by 2.9%, reaching 98.3%. This improvement surpasses recent mainstream algorithms and suggests stronger robustness across various scenarios. The algorithm meets the lightweight requirements for embedded devices in power-line scenarios.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001286893900001  
Indexed Date:  
2024-08-14  
Institution Address:  
Hebei Univ Sci & Technol, Coll Informat Sci & Engn, Shijiazhuang 050018, Peoples R China  
Institution Address:  
Hebei Univ Technol, Coll Elect Engn, Tianjin 300401, Peoples R China  
Institution:  
Institution:

### 387. Cross Dimensional Subspace Attention for human pose estimation

摘要: This paper discusses and studies the importance of Cross Dimensional Subspace Attention for human pose estimation. The Cross Dimensional Subspace Attention module is to realize the interaction between the channel and spatial dimensions of the feature map by dimensional transposition and to learn the individual attention of the feature map subspace, to obtain the global dependence of feature in the feature space and strengthen the ignored semantic and spatial information. This paper focuses on the feature fusion of human pose estimation network. In our method, the feature tensor is rotated and partitioned to obtain the cross-dimension interaction information and learn different mapping of features in subspace, which is conducive to multi-scale feature fusion. On the basis of HRNet (High-Resolution Net) backbone network, our training and testing results on the COCO keypoint detection dataset and the MPII dataset show that our method can effectively improve the performance of human pose estimation.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0141-9382  
eISSN:  
1872-7387  
入藏号:  
WOS:001235203700001  
Indexed Date:  
2024-06-05  
Institution Address:  
Shanghai Univ, Shanghai Inst Adv Commun & Data Sci, Key Lab Specialty Fiber Opt & Opt Access Networks, Sch Commun & Informat Engn,Joint Int Res Lab Speci, Shanghai 200444, Peoples R China  
Institution Address:  
Shenzhen Univ, Coll Informat Engn, Shenzhen 518060, Peoples R China  
Institution:  
Institution:

### 388. Local to non-local: Multi-scale progressive attention network for image

摘要: Image restoration (IR) tasks aim to form a balance between complex textures and spatial details. To this end, the combination of local and non-local attention mechanisms has been well studied in recent years. However, existing local attention-based modules ignore the interaction between channel and spatial attention, while non-local attention operations solely focus on short-range or long-range dependency. To overcome these problems, a novel multi-scale progressive attention network is proposed in this paper, which is termed as MPANet. The proposed MPANet is composed of two parts, a local multi-scale feature extractor and a window -dilation self-attention module. In the local multi-scale feature extractor, a single-scale feature enhancement strategy is designed to model the correlation between channel and spatial dimensions, and a multi-scale feature fusion strategy is applied to further exchange contextual information across all the scales. Furthermore, the window-dilation self-attention is introduced to establish global representation while preserving local details. Experimental results on four IR tasks demonstrate that the proposed MPANet outperforms the state-of-the-art methods in both quantitative results and visual perception.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
4  
ISSN:  
1077-3142  
eISSN:  
1090-235X  
入藏号:  
WOS:001010410200001  
Indexed Date:  
2023-07-01  
Institution Address:  
Tianjin Univ, Sch Elect & Informat Engn, Weijin Rd, Tianjin 300072, Peoples R China  
Institution Address:  
Univ Sci & Technol Beijing, Sch Econ & Management, Beijing, Peoples R China  
Institution:  
Institution:

### 389. InstrumentNet: An integrated model for real-time segmentation of

摘要: In robot-assisted surgery, precise surgical instrument segmentation technology can provide accurate location and pose data for surgeons, helping them perform a series of surgical operations efficiently and safely. However, there are still some interfering factors, such as surgical instruments being covered by tissue, multiple surgical instruments interlacing with each other, and instrument shaking during surgery. To better address these issues, an effective surgical instrument segmentation network called InstrumentNet is proposed, which adopts YOLOv7 as the object detection framework to achieve a real-time detection solution. Specifically, a multiscale feature fusion network is constructed, which aims to avoid problems such as feature redundancy and feature loss and enhance the generalization ability. Furthermore, an adaptive feature-weighted fusion mechanism is introduced to regulate network learning and convergence. Finally, a semantic segmentation head is introduced to integrate the detection and segmentation functions, and a multitask learning loss function is specifically designed to optimize the surgical instrument segmentation performance. The proposed segmentation model is validated on a dataset of intracranial surgical instruments provided by seven experts from Beijing Tiantan Hospital and achieved an mAP score of 93.5 %, Dice score of 82.49 %, and MIoU score of 85.48 %, demonstrating its universality and superiority. The experimental results demonstrate that the proposed model achieves good segmentation performance on surgical instruments compared to other advanced models and can provide a reference for developing intelligent medical robots.  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:001096467800001  
PubMed ID:  
37839219  
Indexed Date:  
2023-11-16  
Institution Address:  
Tianjin Univ Technol, Sch Mech Engn, Tianjin Key Lab Adv Mechatron Syst Design & Intell, Tianjin 300384, Peoples R China  
Institution Address:  
Tianjin Univ Technol, Natl Demonstrat Ctr Expt Mech & Elect Engn Educ, Tianjin, Peoples R China  
Institution Address:  
RIkagaku KENkyusho, Tokyo, Japan  
Institution Address:  
Univ Tokyo, Res Ctr Adv Sci & Technol, Tokyo, Japan  
Institution Address:  
Tianjin Univ Technol, 391 Binshui West Rd, Tianjin 300384, Peoples R China  
Institution:  
Institution:  
Institution:  
RIkagaku KENkyusho  
Institution:  
Institution:

### 390. Multi-Camera System: Imaging Enhancement and Application

摘要: A camera mimics the biological vision system for acquiring natural scenes. Human beings and other animals are mostly equipped with binocular stereo or multi-eye vision systems for high-fidelity scene acquisition, which motivates us to develop a multi-camera system to enhance the imaging capacity. Multi-dimensional and multiscale image/video acquisitions can be realized via heterogeneous sensors and shooting conditions of different cameras, and then leveraging matches across cameras to realize computational multiscale feature fusion for final enhanced reconstruction in respect of scale or dimensionality. In this article, we exemplify the multi-camera system in the applications of wide-field ultrahigh-definition imaging, high spatiotemporal video acquisition, high dynamicrange and low-light imaging enhancement to demonstrate its advantages in improving the imaging capacity, e. g., increasing imaging spatiotemporal resolution, expanding the field of view and extending imaging dynamic range.  
摘要:  
单相机是生物单目视觉的直观模拟,而自然界中的生物主要以双目及多目视觉系统实现场景的精确感知。受此启发,研究人员认为可以应用多相机阵列来增强成像质量。通过不同相机的不同传感器和拍摄条件,实现多维多尺度图像/视频的采集,而后利用相机间的匹配实现不同特征域上的计算融合,从而达到重建场景在尺度或维度上的增强。本文以宽视场超高清成像、时空高速视频采集、高动态范围成像及低照度成像为例,详细介绍了已有多相机系统在成像增强方面的应用,归纳总结多相机系统在提高成像时空分辨率、扩大成像视场、丰富成像动态范围等方面的优势。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7079419  
Indexed Date:  
2021-01-01  
Institution Address:  
南京大学电子科学与工程学院, 南京, 江苏 210023, 中国  
Institution Address:  
School of Electronic Science and Engineering, Nanjing University, Nanjing, Jiangsu 210023, China  
Institution:  
南京大学电子科学与工程学院  
Institution:  
School of Electronic Science and Engineering, Nanjing University

### 391. Dual SIE-FPN: Semantic and Spatial Information Enhancement for

摘要: Feature pyramid network (FPN) can highly improve the performance of object detection by extracting multiscale features. However, current FPN-based methods suffer from intrinsic correlation of local information loss in each feature map, which brings about the semantic information effective transmission problem. In addition, 1 x 1 convolution in lateral connection of FPN may cause spatial information loss. In this article, we propose a novel semantic and spatial information enhancing feature pyramid network (Dual SIE-FPN), which mainly focuses on alleviating multiscale hierarchical feature transmission loss and enhancing the feature representation. Specifically, Dual SIE-FPN contains three modules: Lateral Feature Enhancement (LFE), Global Attention Upsampling (GAU), and Multiple Information Compensation (MIC). LFE is designed to capture deep semantic representation and enhance channel information. GAU is established to make up for spatial information loss caused by upsampling, and transmit the high-level features with the compensatory information to low-level features simultaneously. MIC is designed to work with LFE in parallel to further improve the information loss resulting from 1 x 1 convolution. Experimental results on MS COCO and UAVDT dataset demonstrate that Dual SIE-FPN achieves competitive performance compared to other state-of-the-art FPNs. In addition, our proposed Dual SIE-FPN can be embedded into any multiscale feature extraction-based computer vision tasks to improve the performance.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1551-3203  
eISSN:  
1941-0050  
入藏号:  
WOS:001303387100001  
Indexed Date:  
2024-09-07  
Institution Address:  
Chongqing Univ Posts & Telecommun, Coll Automat, Chongqing 400065, Peoples R China  
Institution Address:  
Inha Univ, Dept Elect & Comp Engn, Incheon 22212, South Korea  
Institution Address:  
Zhejiang Normal Univ, Sch Comp Sci & Technol, Jinhua 321001, Peoples R China  
Institution:  
Institution:  
Institution:

### 392. Optimization Strategy for X-Ray Generation and Countermeasure Fusion of

摘要: During the non-contact flaw detection of a rust-covered bronze mirror,X-ray imaging typically fails to reveal the extent of damage due to the thickness difference between the mirror edge and core.In this study,the X-ray signal from a bronze mirror was used as an input to construct a generative confrontation fusion network.An optimization strategy that enhances the bronze mirror X-ray information fusion was designed to address the reconstruction blur caused by the L\_2 loss and gradient operator,and the expression of multiscale feature details,such as textures and cracks.By utilizing the feature learning process of the L\_(2,1/2) loss regularization generator,the smoothing of the data that was generated using the L\_2 loss was improved;moreover,the Laplacian L\_(tex) pattern loss was defined to strengthen the effect of training network on the extraction of decorations and diseases.Furthermore,a multiscale feature fusion module was added to the training network to improve the quality of the generated information.Thus,considering the experimental comparison involving seven fusion methods,the cross entropy value of the proposed algorithm in two of the five groups is poor.However,the values are optimal in the control data,including entropy,average gradient,spatial frequency,joint entropy,and non-reference feature mutual information.This can effectively reveal the detection information of the bronze mirror during X-ray flaw detection.  
摘要:  
锈蚀覆盖的古铜镜在非接触探伤检测中,因镜缘与镜心厚度各异,X光成像无法呈现完整的病害信息。以古铜镜X光信号为输入,搭建生成对抗融合网络。针对L\_2损失和梯度算子所导致的重构模糊、纹饰和裂痕等多尺度特征细节表达等问题,设计了能够增强古铜镜X光信息融合效果的优化策略。通过添加L\_(2,1/2)损失正则化生成器的特征学习过程,改善L\_2损失生成信息平滑的现象;定义拉普拉斯L\_(tex)纹饰损失,加强训练网络对纹饰和病害的抽取效果;在训练网络中加入多尺度特征融合模块,提高细节信息生成质量。通过与7种融合方法进行实验对比,所提算法在5组对照数据中仅2组的交叉熵值略差,其余信息熵、平均梯度、空间频率、联合熵和非参考特征互信息值均取得最优,可有效呈现古铜镜X光探伤检测信息。  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7404613  
Indexed Date:  
2023-06-02  
Institution Address:  
西安建筑科技大学信息与控制工程学院, 西安, 陕西 710055, 中国  
Institution Address:  
陕西省文物保护研究院, 西安, 陕西 710075, 中国  
Institution Address:  
College of Information and Control Engineering,Xi'an University of Architecture and Technology, Xi'an, Shaanxi 710055, China  
Institution Address:  
Shaanxi Institute for the Preservation of Cultural Heritage, Xi'an, Shaanxi 710075, China  
Institution:  
西安建筑科技大学信息与控制工程学院  
Institution:  
陕西省文物保护研究院  
Institution:  
College of Information and Control Engineering,Xi'an University of Architecture and Technology  
Institution:  
Shaanxi Institute for the Preservation of Cultural Heritage

### 393. Dual-Branch Low-Light Image Enhancement Combined with Dense Wavelet

摘要: A dual-branch image enhancement method combining dense wavelet transform is proposed to solve the problems of low brightness, high noise, and color distortion in low-light images. Firstly, dense wavelet networks are used for multi-scale feature information fusion to reduce information loss and provide denoising capability. Then, the global attention module and feature extraction module are embedded in the multi-scale feature fusion to fully extract global and local features. Finally, the effect of low-light images is enhanced by color enhancement and detail reconstruction with a dualbranch structure. In addition, a new joint loss function is introduced to guide the network training from multiple aspects to enhance its performance. The experimental results show that the proposed method effectively improves the brightness of low-light images, suppresses image noise, and obtains richer details and color information. The enhanced images are clearer and more natural, and the peak signal-to-noise ratio and structural similarity have significant advantages over the mainstream methods.  
摘要:  
针对低照度图像存在低亮度、高噪声、色彩失真等问题,提出了一种结合稠密小波变换的双分支低照度图像增强方法。采用稠密小波网络进行多尺度特征信息融合,在减少信息丢失的同时使网络具有一定的去噪能力。在多尺度特征融合中嵌入全局注意力模块和特征提取模块,充分提取全局和局部特征。通过双分支结构对图像进行色彩增强和细节重建,使得低照度图像具有较好的增强效果。引入了新的联合损失函数从多方面指导网络训练,以增强网络性能。将所提方法与主流方法相比较,实验结果充分表明,所提方法有效提高了低照度图像的亮度,抑制了图像噪声,并取得了更丰富的细节和色彩信息,得到的增强图像更清晰自然,在峰值信噪比和结构相似度等图像质量客观评价指标方面也具有显著的优势。  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1002-8331  
入藏号:  
CSCD:7664194  
Indexed Date:  
2024-05-17  
Institution Address:  
中国计量大学,信息工程学院;;中国计量大学, ;;浙江省电磁波信息技术与计量检测重点实验室, 杭州;;杭州, ;; 310018;;310018  
Institution Address:  
College of Information Engineering, China Jiliang University;;China Jiliang University, ;;Key Laboratory of Electromagnetic Wave Information Technology and Metrology of Zhejiang Province, Hangzhou;;Hangzhou, ;; 310018;;310018  
Institution:  
中国计量大学,信息工程学院;;中国计量大学  
Institution:  
College of Information Engineering, China Jiliang University;;China Jiliang University

### 394. HRU-Net: High-Resolution Remote Sensing Image Road Extraction Based on

摘要: Road extraction from high-resolution satellite images has become a significant focus in the field of remote sensing image analysis. However, factors such as shadow occlusion and spectral confusion hinder the accuracy and consistency of road extraction in satellite images. To overcome these challenges, this paper presents a multi-scale fusion-based road extraction framework, HRU-Net, which exploits the various scales and resolutions of image features generated during the encoding and decoding processes. First, during the encoding phase, we develop a multi-scale feature fusion module with upsampling capabilities (UMR module) to capture fine details, enhancing shadowed areas and road boundaries. Next, in the decoding phase, we design a multi-feature fusion module (MPF module) to obtain multi-scale spatial information, enabling better differentiation between roads and objects with similar spectral characteristics. The network simultaneously integrates multi-scale feature information during the downsampling process, producing high-resolution feature maps through progressive cross-layer connections, thereby enabling more effective high-resolution prediction tasks. We conduct comparative experiments and quantitative evaluations of the proposed HRU-Net framework against existing algorithms (U-Net, ResNet, DeepLabV3, ResUnet, HRNet) using the Massachusetts Road Dataset. On this basis, this paper selects three network models (U-Net, HRNet, and HRU-Net) to conduct comparative experiments and quantitative evaluations on the DeepGlobe Road Dataset. The experimental results demonstrate that the HRU-Net framework outperforms its counterparts in terms of accuracy and mean intersection over union. In summary, the HRU-Net model proposed in this paper skillfully exploits information from different resolution feature maps, effectively addressing the challenges of discontinuous road extraction and reduced accuracy caused by shadow occlusion and spectral confusion factors. In complex satellite image scenarios, the model accurately extracts comprehensive road regions.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
2076-3417  
入藏号:  
WOS:001037972400001  
Indexed Date:  
2023-08-08  
Institution Address:  
Guilin Univ Technol, Coll Geomat & Geoinformat, Guilin 541004, Peoples R China  
Institution:

### 395. PDR-Net: Progressive depth reconstruction network for color guided depth

摘要: Low spatial resolution is a common problem for depth maps captured by low-cost consumer depth cameras. Depth map super-resolution (DMSR) can improve the quality of depth maps, but it is an ill-posed problem with many challenges. This paper proposes a progressive depth reconstruction network (PDRNet) to further enhance the performance of DMSR. Specifically, we design an adaptive feature recombination module to recombine depth and color guidance features. We generate sufficient information from the recombined features with the proposed multi-scale feature fusion module, in which multi-scale feature distillation and joint attention mechanism are employed. We learn high frequency compensations for each up-interpolating and reconstruct corresponding high resolution depth maps in the proposed progressive depth reconstruction module. Experimental results with benchmark datasets verified the proposed method's superiority over the state-of-the-art DMSR methods. (c) 2022 Elsevier B.V. All rights reserved.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
14  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000761683700007  
Indexed Date:  
2022-03-18  
Institution Address:  
Hebei Univ Technol, Sch Mech Engn, Tianjin 300130, Peoples R China  
Institution Address:  
Tangshan Coll, Sch Intelligence & Informat Engn, Tangshan 063000, Peoples R China  
Institution:  
Institution:  
Tangshan Coll

### 396. SDD-Net: A Steel Surface Defect Detection Method Based on Contextual

摘要: Addressing the current industrial methods for surface defect detection, which suffer from issues such as low detection efficiency, elevated rates of false positives, and inadequate real-time capabilities, this paper proposes an high-precision industrial defect detection network. Firstly, this paper proposes a lightweight feature extraction network, which ensures real-time model detection under the premise of fully extracting defect features. Secondly, in order to solve the problem that tiny targets in industrial datasets have fuzzy texture and contain few features that are difficult to be detected, a Context Enhancement Module (CEM) is proposed, which effectively complements the contextual information of the small targets and performs multi-scale fusion to enhance the semantic information representation. Meanwhile, a Feature Enhancement Module (FEM) is designed at the end of the backbone network to optimise the feature information, effectively capture the global feature information and local feature information, and the Hybrid Attention Module (HAM) designed in this paper is introduced to extract the important information and weaken the irrelevant information in the feature map. Finally, a Dense cross-layer Feature Pyramid Network (DFPN) is proposed, which fully integrates the semantic and fine-grained features extracted from the backbone. Improve the detection of targets with significant scale changes, and adds a feature refinement module to suppress the conflicting information and reduce the semantic discrepancies before passing the learned multi-scale feature information into the prediction layer. The surface of the experimental results,the mAP@.5 was obtained at 94.3%, 98.6% and 98.4% on the steel, PCB and aluminium surface defect datasets.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001378949400032  
Indexed Date:  
2024-12-28  
Institution Address:  
Heilongjiang Univ Sci & Technol, Sch Safety Engn, Harbin 150022, Peoples R China  
Institution Address:  
Shaanxi Univ Sci & Technol, Ulster Coll, Xian 710021, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Sch Modern Posts, Chongqing 400065, Peoples R China  
Institution Address:  
Qinghai Minzu Univ, Sch Math & Stat, Xining 810007, Qinghai, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Qinghai Minzu Univ

### 397. SOCR-YOLO : Small Objects Detection Algorithm in Medical Images

摘要: In the field of medical image analysis, object detection plays a crucial role by providing interpretable diagnostic information to healthcare professionals. Although current object detection models have achieved remarkable success in conventional images, their performance in detecting abnormalities in medical images has not been as satisfactory. This is primarily due to the complexity of anatomical structures in medical images, and the fact that some lesions may have subtle features, particularly in the case of early-stage, small-scale abnormalities. To address this challenge, we introduce SOCR-YOLO, a novel lesion detection model with online convolutional reparameterization based on channel shuffling. First, it employs the SOCR (Shuffled Channel with Online Convolutional Re-parameterization) module to establish a connection between feature concatenation and computational efficiency, aiming to extract more comprehensive information while reducing time consumption. Second, it incorporates the Bi-FPN structure to achieve multiscale feature fusion. Lastly, the loss function has been optimized to improve the model training process. We evaluated two datasets, chest x-ray (Vindr-CXR) and brain tumor (Br35H), provided by the Kaggle competition. Experimental results show that the proposed method has outperformed several state-of-the-art models, including YOLOv8, YOLO-NAS, and RT-DETR, in both speed and accuracy. Notably, in the context of chest x-ray anomaly detection, SOCR-YOLO exhibits a 1.8% enhancement in accuracy over YOLOv8 while simultaneously reducing floating-point operations by 26.3%. Additionally, a similar 1.8% improvement in accuracy is observed in the detection of brain tumors. The results indicate the superior ability of our model to detect multiscale variations and small lesions.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0899-9457  
eISSN:  
1098-1098  
入藏号:  
WOS:001252291100001  
Indexed Date:  
2024-06-30  
Institution Address:  
Zhejiang Sci Tech Univ, Sch Comp Sci & Technol, Hangzhou, Peoples R China  
Institution Address:  
903rd Hosp Chinese Peoples Liberat Army, Dept Neurosurg, Hangzhou, Peoples R China  
Institution Address:  
Imperial Coll London, Natl Heart & Lung Inst, London, England  
Institution Address:  
Imperial Coll London, Bioengn Dept & Imperial X, London, England  
Institution Address:  
Royal Brompton Hosp, Cardiovasc Res Ctr, London, England  
Institution Address:  
Kings Coll London, Sch Biomed Engn & Imaging Sci, London, England  
Institution:  
Institution:  
903rd Hosp Chinese Peoples Liberat Army  
Institution:  
Institution:  
Institution:  
Institution:

### 398. Attention mechanism and multi-scale feature fusion network for ALS point

摘要: Airborne Laser Scanning(ALS)point cloud classification is essential to extract geoinformation,but the uneven spatial distribution and scale variations between different categories bring challenges to the fine classification of point cloud data.In this paper,an attention mechanism and multiscale feature fusion deep learing network(AMMSF-Net)for ALS point cloud classification was proposed. In the network a local spatial position attention layer was used to learn local contextual features;and an attention skip connection scheme was added to dynamic fusion the corresponding features among the encoder and decoder,which can retain detail features and contextual information.The multi-scale feature in the decoder fusion module obtained the final semantic probability map by concatenating the features at different scales into MLP(Multilayer Perceptron)and CML(Conditional Markov Layer),which can achieve the correlation of the feature maps between different scales and different levels,and can enhance the expression of targets at different scales.In two datasets experiment,AMMSF-Net achieveed 83.8%overall accuracy and 70.4%average F1score in the Vaihingen dataset,and 95.4% overall accuracy and 88.5% average F1score in the DFC3D dataset.Compared with other popular method,AMMSF-Net got higher classification accuracy in both datasets, which shows that AMMSF-Net can distinguish ground objects in point cloud effectively.  
摘要:  
针对机载激光点云数据中存在空间分布不均匀和地物尺度不一的问题,提出了一种融合注意力机制和多尺度特征的机载激光点云分类深度学习方法(AMMSF-Net)。该方法建立了局部空间位置注意力层学习局部邻域上下文特征,增加注意力跳连机制将解码器和编码器中的特征进行动态融合并有效保留细节信息;解码器中的多尺度特征融合通过将不同尺度的特征进行级联输入到多层感知机和条件马尔可夫层得到最后的语义概率图,实现了不同尺度与不同层级特征图之间的相关,增强不同尺度目标的表达能力。在Vaihingen数据集中AMMSF-Net取得83.8%的总体精度和70.4%平均F1分数,在DFC3D数据集取得了95.4%总体精度和88.5%平均F1分数,对比其他模型该方法在两个数据集都取得了更好的精度,这表明AMMSF-Net能有效提高点云地物类别区分的能力。  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1009-2307  
入藏号:  
CSCD:7384959  
Indexed Date:  
2023-05-12  
Institution Address:  
西安科技大学测绘科学与技术学院, 西安, 陕西 710054, 中国  
Institution Address:  
天津市勘察设计院集团有限公司, 天津 300391, 中国  
Institution Address:  
College of Geomatics, Xi'an University of Science and Technology, Xi'an, Shaanxi 710054, China  
Institution Address:  
Tianjin Institute of Geotechnical Investigation Surveying, Tianjin 300391, China  
Institution:  
西安科技大学测绘科学与技术学院  
Institution:  
天津市勘察设计院集团有限公司  
Institution:  
College of Geomatics, Xi'an University of Science and Technology  
Institution:  
Tianjin Institute of Geotechnical Investigation Surveying

### 399. Human pose estimation with gated multi-scale feature fusion and spatial

摘要: Although human pose estimation has achieved great success, the ambiguity of joint prediction has not been well resolved, especially in complex situations (crowded scenes, occlusions, and unnormal poses). We think that is caused by the noisy information introduced by combining multi-level features by simply adding features at each position. To alleviate this problem, we propose a new structure of gated multi-scale feature fusion (GMSFF). This module aims to selectively import high-level features to make up for the missing semantic information of low-resolution feature maps. Inspired by the prior knowledge that the position information of joints can refer to each other, we propose a new fine-tuning strategy for pose estimation-spatial mutual information complementary module (SMICM). It can assist the model in better adjusting the current joint's position by capturing the information contained in other joints and only adds a little computational cost. We evaluated our proposed method on four datasets: MPII Human Pose Dataset (MPII), COCO keypoint detection Dataset (COCO), Occluded Human Dataset (OCHuman), and CrowdPose Dataset. The experimental results show that with the deepening of the occlusion and crowding level of the datasets, the improvement becomes more and more obvious. In particular, a performance improvement of 2.2 AP was obtained on the OCHuman dataset. In addition, our modules are plug-and-play.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:000715675100001  
Indexed Date:  
2021-11-17  
Institution Address:  
Tianjin Univ, Sch Microelect, Tianjin, Peoples R China  
Institution Address:  
Tianjin Int Joint Res Ctr Internet Things, Tianjin, Peoples R China  
Institution Address:  
Tianjin Univ, Tianjin Key Lab Imaging & Sensing Microelect Tech, Tianjin, Peoples R China  
Institution:  
Institution:  
Tianjin Int Joint Res Ctr Internet Things  
Institution:

### 400. Insulator Semantic Segmentation in Aerial Images Based on Multiscale

摘要: As one of the important components in the transmission line, the insulator is related to the safe and reliable operation of the entire transmission line. Aerial images are characterized by complex backgrounds, multiple pseudotargets, and low signal-to-noise ratios. Rapid and accurate localization of insulators in aerial images is a critical and challenging task in automatic inspection of transmission lines. Most insulator localization methods suffer from the loss of target edge detail information and large amount of model parameters. To solve these problems, this paper adopts an Encoder-Decoder architecture, called ED-Net, to realize end-to-end intelligent and accurate identification of insulators in aerial images. Firstly, Initial Module and CA-Bottleneck which are used to extract features from images to generate finer feature maps are proposed in the Encoder path. Meanwhile, global average pooling is used to preserve the maximum receptive field. Secondly, in the Decoder path, Refinement Boundary Module and Asymmetric Convolution Module are given to perform boundary optimization on the feature map, which are generated by the Encoder path. Finally, the Attention Feature Fusion Module is introduced into the Decoder path to combine high-level features with low-level features better and reduce the gap between features of different levels. The proposed model architecture keeps a suitable balance between the model parameters and insulator segmentation performance on insulator test datasets. Specifically, for a 512 x 512 input image, 95.12% mean intersection over union is achieved on the insulator test datasets with different environments and model parameters size being only 13.61 M. Compared with the current state-of-the-art semantic segmentation methods, the results show that the proposed method has higher efficient and accuracy.  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1076-2787  
eISSN:  
1099-0526  
入藏号:  
WOS:000826055600001  
Indexed Date:  
2022-07-22  
Institution Address:  
Kunming Univ Sci & Technol, Fac Mech & Elect Engn, Kunming 650500, Yunnan, Peoples R China  
Institution:

### 401. Mini-infrared thermal imaging system image denoising with multi-head

摘要: The mini-infrared thermal imaging system (MITIS) has been widely applied in outdoor photography, military observation, medical auxiliary diagnosis, and so on. However, due to the limitations of infrared sensor performance, imaging system size, and environment interference, the infrared images captured by the current MITIS generally encounter noise interference and low contrast issues. To solve these issues, we propose a multi-head feature fusion and detail enhancement network, namely MdNet, for complex features (texture and edges) retaining and denoising of infrared images obtained from the MITIS. Specifically, MdNet consists of a multi-head feature fusion block (MFFB), detail enhancement block (DEB), and reconstruction convolution. The MFFB is designed to extract and fuse multi-scale feature information, including multi-scale and feature fusion blocks. The DEB is developed to enhance the noise feature and reduce the loss of detailed information by three multi-level local feature attention blocks and residual learning. Quantitative and qualitative experimental results demonstrate that the MdNet achieves competitive performance in removing synthetic and natural noises compared with other state-of-the-art infrared image denoising methods and provides an effective solution for MITIS image denoising.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
0030-3992  
eISSN:  
1879-2545  
入藏号:  
WOS:001253737300001  
Indexed Date:  
2024-06-30  
Institution Address:  
Guangdong Univ Technol, Sch Automat, Guangdong Prov Key Lab Cyber Phys Syst, Guangzhou 510006, Peoples R China  
Institution Address:  
Guangdong Univ Technol, Sch Comp, Guangzhou 510006, Peoples R China  
Institution Address:  
Guangdong Univ Technol, Sch Chem Engn & Light Ind, Guangzhou 510006, Peoples R China  
Institution:  
Institution:  
Institution:

### 402. YOLOv5s-SA: Light-Weighted and Improved YOLOv5s for Sperm Detection

摘要: Sperm detection performance is particularly critical for sperm motility tracking. However, there are a large number of non-sperm objects, sperm occlusion and poorly detailed texture features in semen images, which directly affect the accuracy of sperm detection. To solve the problem of false detection and missed detection in sperm detection, a multi-sperm target detection model, Yolov5s-SA, with an SA attention mechanism is proposed based on the YOLOv5s algorithm. Firstly, a depthwise, separable convolution structure is used to replace the partial convolution of the backbone network, which can ensure stable precision and reduce the number of model parameters. Secondly, a new multi-scale feature fusion module is designed to enhance the perception of feature information to supplement the positional information and high-resolution of the deep feature map. Finally, the SA attention mechanism is integrated into the neck network before the output of the feature map to enhance the correlation between the feature map channels and improve the fine-grained feature fusion ability of YOLOv5s. Experimental results show that compared with various YOLO algorithms, the proposed algorithm improves the detection accuracy and speed to a certain extent. Compared with the YOLOv3, YOLOv3-spp, YOLOv5s and YOLOv5m models, the average accuracy increases by 18.1%, 15.2%, 6.9% and 1.9%, respectively. It can effectively reduce the missed detection of occluded sperm and achieve lightweight and efficient multi-sperm target detection.  
Web of Science 核心合集中的 "被引频次":  
8  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
eISSN:  
2075-4418  
入藏号:  
WOS:000955442100001  
PubMed ID:  
36980408  
Indexed Date:  
2023-04-09  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Elect Engn, Beijing 100876, Peoples R China  
Institution Address:  
SAS Med Technol Beijing Co Ltd, Beijing 102200, Peoples R China  
Institution:  
Institution:  
SAS Med Technol Beijing Co Ltd

### 403. Enhancing Medical Image Classification With Context Modulated Attention

摘要: This research proposes a multi-stage feature fusion network (MSFF) for medical image classification. In view of the problems existing in medical images, such as noise, diversity, and similarity among different classes, MSFF enhances the global context perception in the window partitioning framework through Context Modulation Attention (CMA). Meanwhile, it extracts fine-grained local information via the multi-stage Contextual Information Refinement (CIR) module and gradually fuses multi-stage local and global features to generate richer semantic representations. The experimental results demonstrate that MSFF significantly outperforms existing methods in multiple performance metrics (including accuracy, precision, recall, F1-score, Matthews Correlation Coefficient (MCC), Kappa coefficient, Area Under the Curve (AUC), balanced accuracy, and geometric mean) on four datasets (Endoscopic Bladder Tissue, Kvasir, SARS-COV-2 Ct-Scan, and Thyroid Nodule), showing its excellent performance in the task of medical image classification.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001410337900030  
Indexed Date:  
2025-02-09  
Institution Address:  
China West Normal Univ, Sch Comp Sci, Nanchong 637009, Sichuan, Peoples R China  
Institution Address:  
China West Normal Univ, Inst Artificial Intelligence, Nanchong 637000, Sichuan, Peoples R China  
Institution Address:  
Panzhihua Univ, Sch Math & Comp Sci, Panzhihua 617000, Sichuan, Peoples R China  
Institution Address:  
Nanjiang TCM Hosp, Dept Radiol, Bazhong 636600, Sichuan, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Nanjiang TCM Hosp

### 404. Surface Defect Detection for Aerospace Aluminum Profiles with Attention

摘要: A YOLOv5 aluminum profile defect detection algorithm that integrates attention and multi-scale features is proposed in this paper to address the issues of the low detection accuracy, high false detection rates, and high missed detection rates that are caused by the large-scale variation of surface defects, inconspicuous small defect characteristics, and a lack of concentrated feature information in defect areas. Firstly, an improved CBAM (Channel-Wise Attention Module) convolutional attention module is employed, which effectively focuses on the feature information of defect areas in the aluminum defect dataset with only a small amount of spatial dimension. Secondly, a bidirectional weighted feature fusion network is utilized, incorporating a multi-scale feature fusion network with skip connections to aggregate various high-resolution features, thus enriching the semantic expression of features. Then, new size feature maps that have not been fused are introduced into the detection layer network to improve the detection effect of small target defects. Experimental results indicate that an average detection accuracy (mAP) of 82.6% was achieved by the improved YOLOv5 algorithm on the aluminum surface defect dataset. An improvement of 6.2% over the previous version was observed. The current defect detection requirements of aluminum profile production sites are met by this enhanced algorithm.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
被引频次合计:  
6  
eISSN:  
2079-9292  
入藏号:  
WOS:001277040700001  
Indexed Date:  
2024-08-02  
Institution Address:  
Shaanxi Univ Sci & Technol, Sch Elect & Control Engn, Xian 710016, Peoples R China  
Institution:

### 405. Multi-Scale Feature Fusion and Context-Enhanced Spatial Sparse

摘要: Accurate and efficient object detection in UAV images is a challenging task due to the diversity of target scales and the massive number of small targets. This study investigates the enhancement in the detection head using sparse convolution, demonstrating its effectiveness in achieving an optimal balance between accuracy and efficiency. Nevertheless, the sparse convolution method encounters challenges related to the inadequate incorporation of global contextual information and exhibits network inflexibility attributable to its fixed mask ratios. To address the above issues, the MFFCESSC-SSD, a novel single-shot detector (SSD) with multi-scale feature fusion and context-enhanced spatial sparse convolution, is proposed in this paper. First, a global context-enhanced group normalization (CE-GN) layer is developed to address the issue of information loss resulting from the convolution process applied exclusively to the masked region. Subsequently, a dynamic masking strategy is designed to determine the optimal mask ratios, thereby ensuring compact foreground coverage that enhances both accuracy and efficiency. Experiments on two datasets (i.e., VisDrone and ARH2000; the latter dataset was created by the researchers) demonstrate that the MFFCESSC-SSD remarkably outperforms the performance of the SSD and numerous conventional object detection algorithms in terms of accuracy and efficiency.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001403990900001  
Indexed Date:  
2025-01-30  
Institution Address:  
Inner Mongolia Normal Univ, Coll Comp Sci & Technol, Hohhot 010022, Peoples R China  
Institution Address:  
Inner Mongolia Agr Univ, Coll Mech & Elect Engn, Hohhot 010010, Peoples R China  
Institution:  
Institution:

### 406. AFRNet: adaptive feature refinement network

摘要: In the domain of computer vision, object detection is a fundamental task, aimed at accurately identifying and localizing objects of various sizes within images. While existing models such as You Only Look Once, Adaptive Training Sample Selection, and Task-aligned One-stage Object Detection have made breakthroughs in this field, they still exhibit deficiencies in information fusion within their neck structure. To overcome these limitations, we have designed an innovative model architecture known as Adaptive Feature Refinement Network (AFRNet). The model, on one hand, discards the conventional Feature Pyramid Network structure and designs a novel neck structure that incorporates the structures of Scale Sequence Feature Fusion (SSFF) model and the Gather-and-Distribute (GD) mechanism. Through experimentation, it has been demonstrated that the SSFF method can further enhance the multi-scale feature fusion of the GD mechanism, thereby improving the performance of the target detection task. On the other hand, to address the constraints of existing models in simulating geometric transformations, We have designed an advanced variable convolution structure called Attentive Deformable ConvNet. This structure integrates an improved attention mechanism, which allows for more precise capture of key features in images. Extensive experiments conducted on the MS-COCO dataset have validated the effectiveness of our model. In single-model, single-scale testing, our model achieved an Average Precision (AP) of 51.8%, a result that underscores a significant enhancement in object detection performance and confirms the efficacy of our model.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001272292500002  
Indexed Date:  
2024-07-26  
Institution Address:  
Northwest Normal Univ, Coll Comp Sci & Engn, Lanzhou 730070, Gansu, Peoples R China  
Institution:

### 407. Low-Light Salient Object Detection by Learning to Highlight the

摘要: Previous methods in salient object detection (SOD) mainly focused on favorable illumination circumstances while neglecting the performance in low-light condition, which significantly impedes the development of related down-stream tasks. In this work, considering that it is impractical to annotate the large-scale labels for this task, we present a framework (HDNet) to detect the salient objects in low-light images with the synthetic images. Our HDNet consists of a foreground highlight sub-network (HNet) and an appearance-aware detection sub-network (DNet), both of which can be learned jointly in an end-to-end manner. Specifically, to highlight the foreground objects, we design the HNet to estimate the parameters to adjust the dynamic range for each pixel adaptively, which can be trained via the weak supervision signals of the salient object labels. In addition, we design a simple detection network (DNet) with a contextual feature fusion module and a multi-scale feature refine module for detailed feature fusion and refinement. Furthermore, we contribute the first annotated dataset for salient object detection in low-light images (SOD-LL), including 6,000 labeled synthetic images (SOD-LLS) and 2,000 labeled real images (SOD-LLR). Experimental results on SOD-LL and other low-light videos in the wild demonstrate the effectiveness and generalization ability of our method. Our dataset and code are available at https://github.com/Ylinyuan/HDNet.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:001327614800077  
Indexed Date:  
2024-11-15  
Institution Address:  
Hunan Normal Univ, Coll Engn & Design, Changsha 410081, Peoples R China  
Institution Address:  
Western Univ, Dept Elect & Comp Engn, London, ON N6A 3K7, Canada  
Institution:  
Institution:

### 408. MDAU-Net: A Liver and Liver Tumor Segmentation Method Combining an

摘要: In recent years, U-Net and its extended variants have made remarkable progress in the realm of liver and liver tumor segmentation. However, the limitations of single-path convolutional operations have hindered the full exploitation of valuable features and restricted their mobility within networks. Moreover, the semantic gap between shallow and deep features proves that a simplistic shortcut is not enough. To address these issues and realize automatic liver and tumor area segmentation in CT images, we introduced the multi-scale feature fusion with dense connections and an attention mechanism segmentation method (MDAU-Net). This network leverages the multi-head attention (MHA) mechanism and multi-scale feature fusion. First, we introduced a double-flow linear pooling enhancement unit to optimize the fusion of deep and shallow features while mitigating the semantic gap between them. Subsequently, we proposed a cascaded adaptive feature extraction unit, combining attention mechanisms with a series of dense connections to capture valuable information and encourage feature reuse. Additionally, we designed a cross-level information interaction mechanism utilizing bidirectional residual connections to address the issue of forgetting a priori knowledge during training. Finally, we assessed MDAU-Net's performance on the LiTS and SLiver07 datasets. The experimental results demonstrated that MDAU-Net is well-suited for liver and tumor segmentation tasks, outperforming existing widely used methods in terms of robustness and accuracy.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001072545500001  
Indexed Date:  
2023-11-03  
Institution Address:  
North Minzu Univ, Sch Comp Sci & Engn, Yinchuan 750021, Peoples R China  
Institution Address:  
North Minzu Univ, Key Lab Images & Graph Intelligent Proc Natl Ethn, Yinchuan 750021, Peoples R China  
Institution Address:  
North Minzu Univ, Sch Math & Informat Sci, Yinchuan 750021, Peoples R China  
Institution:  
Institution:  
Institution:

### 409. Few-Shot Classification with Multiscale Feature Fusion for Clinical Skin

摘要: Introduction: Skin disease is one of the most common diseases and can affect people of all ages and races. However, the diagnosis of skin diseases via observation is a highly challenging task for both doctors and patients, and would benefit from the use of an intelligent system. Building a large benchmark with professional dermatologists is resource-intensive, and we believe that few-shot learning (FSL) methods would be helpful in solving the problem of annotated data scarcity. In this paper, we propose CDD-Net (Context Feature Fusion and Dual Attention Dermatology Net), a plug-in module for FSL clinical skin disease classification. Methods: Current FSL methods used in skin disease classification are limited to nonuniversal approaches and few disease classes. Our CDD-Net has a flexible structure, including a context feature-fusion module and dual-attention module to extract discriminating texture feature and emphasize contributive regions and channels. The context feature-fusion module localizes discriminatory texture details of skin lesions by integrating features from different layers, while the dual-attention module highlights discriminative regions via channel-wise and pixel-wise depictions based on weight vectors and restrains the contributions of irrelevant areas. We also present Derm104, a new clinical skin disease data benchmark that has significant coverage of rare diseases and reliable annotation between primary species and subspecies for better validation of our approach. Results: Our experiments validated the versatility of CDD-Net for different FSL methods and achieved an improvement in accuracy of up to 9.14 percentage points compared with the vanilla network, which can be considered state of the art. The ablation study also showed that the dual-attention module and context feature-fusion module worked efficiently in CDD-Net.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1178-7015  
入藏号:  
WOS:001382094600001  
PubMed ID:  
38737944  
Indexed Date:  
2024-12-28  
Institution Address:  
Sichuan Univ, Coll Biomed Engn, Chengdu 610065, Peoples R China  
Institution:

### 410. DMC-UNet-Based Segmentation of Lung Nodules

摘要: The accurate and rapid segmentation of different categories of lung nodules is of great importance for the diagnosis of early stage lung cancer and to assist physicians in the diagnosis and treatment of the disease. In the segmentation process, there are various types of lung nodules with different shape characteristics and occupying small volumes, so the process of segmenting lung nodules out is challenging. The DMC-UNet network proposed in this paper is an improved network based on UNet. The DMC-UNet network combines a lightweight residual structure, multiscale feature upsampling fusion and X/Y Channel Attention Module and Coordinate Attention (CCA) attention mechanism. The overall framework of the network firstly replaces the convolutional units of U-Net with residual units, and replaces the traditional convolution in the residual units with Depthwise Separable Convolution (DSC) to reduce the number of parameters and computation of the model and improve the efficiency of model training and prediction, and secondly replaces the transposed convolution and PixelShuffle in the upsampling process of U-Net with parallel direction fusion to replace the transposed convolution used in the original U-Net, which can enable the model to better capture information at different scales, and the addition of a multiscale feature fusion module before PixelShuffle improves the traditional Efficient Sub-Pixel Convolutional Neural(ESPCN) model, which aims to expand the perceptual field, and finally, the addition of a CCA attention mechanism after the upsampling fusion can better recover spatial information. It is shown by experiments that the IoU and F1-score of DMC-UNet are 65.52%+/- 0.71% and 76.02%+/- 0.63%, respectively, on the lung nodules provided by the Department of Medical Imaging of the Fourth Affiliated Hospital of Guangxi Medical University (FAHGMU), and the absolute gains of IoU and F1-score compared with U-Net are 2.78% and 2.91% on the Lung Image Database Consortium(LIDC) public dataset, and 83.36% and 89.92% on the IoU and F1-score, respectively, with a gain of 1.37% and 0.73% on the IoU and F1-score, respectively, compared to the U-Net.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001085254800001  
Indexed Date:  
2023-11-01  
Institution Address:  
Guangxi Univ Sci & Technol, Sch Automat, Liuzhou, Peoples R China  
Institution Address:  
Guangxi Med Univ, Affiliated Hosp 4, Liuzhou 545000, Peoples R China  
Institution:  
Institution:

### 411. Pedestrian detection algorithm based on multi-scale feature extraction

摘要: Despite rapid development in anchor-free pedestrian detection algorithms, an optimistic trade-off between detection accuracy and efficiency is still far from being achieved. In this study, we proposed a new pedestrian detection algorithm based on multi-scale feature extraction and attention feature fusion, which is called MSAF-Net. Firstly, we designed a multi-scale dilate residual module to expand receptive fields while maintaining feature map sizes and improving the spatial sensitivity of information features. Secondly, through a joint attention feature fusion mechanism, the interaction of channel and space joint information features was captured. The context and initial features were integrated into strengthen extracted image features. Finally, a channel attention guidance mask branch was added to a detector to locate pedestrian position information accurately, and to improve the model's robustness. The experimental results demonstrated that our algorithm achieved satisfactory results on two well established datasets, i.e., CityPersons and Caltech. For CityPersons, the values of MR-2 were 9.04% and 40.41% under reasonable and heavy occlusion conditions, which are 0.36% and 3.89% better than the suboptimal comparison detection methods, proving the effectiveness and advancement of the proposed algorithm in this study. (C) 2021 Elsevier Inc. All rights reserved.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
15  
被引频次合计:  
15  
ISSN:  
1051-2004  
eISSN:  
1095-4333  
入藏号:  
WOS:000730121600001  
Indexed Date:  
2021-06-04  
Institution Address:  
Shandong Normal Univ, Sch Phys & Elect Sci, Jinan 250358, Shandong, Peoples R China  
Institution:

### 412. Small object detection based on hierarchical attention mechanism and

摘要: The ability of modern detectors to detect small targets is still an unresolved topic compared to their capability of detecting medium and large targets in the field of object detection. Accurately detecting and identifying small objects in the real-world scenario suffer from sub-optimal performance due to various factors such as small target size, complex background, variability in illumination, occlusions, and target distortion. Here, a small object detection method for complex traffic scenarios named deformable local and global attention (DLGADet) is proposed, which seamlessly merges the ability of hierarchical attention mechanisms (HAMs) with the versatility of deformable multi-scale feature fusion, effectively improving recognition and detection performance. First, DLGADet introduces the combination of multi-scale separable detection and multi-scale feature fusion mechanism to obtain richer contextual information for feature fusion while solving the misalignment problem of classification and localisation tasks. Second, a deformation feature extraction module (DFEM) is designed to address the deformation of objects. Finally, a HAM combining global and local attention mechanisms is designed to obtain discriminative features from complex backgrounds. Extensive experiments on three datasets demonstrate the effectiveness of the proposed methods. Code is available at  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1751-9659  
eISSN:  
1751-9667  
入藏号:  
WOS:001051841500001  
Indexed Date:  
2023-08-31  
Institution Address:  
Henan Univ, Sch Software, Kaifeng, Henan, Peoples R China  
Institution Address:  
State Grid Henan Prov Elect Power Co, Econ & Tech Res Inst, Zhengzhou, Henan, Peoples R China  
Institution Address:  
Henan Univ, Sch Econ, Kaifeng, Henan, Peoples R China  
Institution Address:  
Shanghai Inst Microsyst & Informat Technol, Elect & Comp Engn, Shanghai, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 413. A new anti-vibration hammer rust detection algorithm based on improved

摘要: Anti-vibration hammer corrosion seriously endangers the safe operation of transmission lines. To better improve the detection effect of anti-vibration hammer corrosion on transmission lines under complex background, in this paper we propose an anti-vibration hammer corrosion detection algorithm based on improved YOLOv7. Firstly, a dataset was established for anti-vibration hammer corrosion detection. After that, a color space transformation was carried out through an HSV color model to highlight the corrosion features. BiFPN multi-scale feature fusion was then introduced to fully utilize the feature maps at different scales, so as to get a weighted feature fusion of weights. Finally, a GSConv module was added to the YOLOv7 network to avoid spatial information loss in the transmission process as far as possible. It can accelerate the inference speed and increase the model accuracy as well. The test results show that the improved algorithm improves the detection accuracy by 1.6% and inference speed by 17% on average compared to the original network. (c) 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under theCCBY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).  
会议名称:  
3rd International Conference on Power and Electrical Engineering  
会议日期:  
DEC 29-31, 2022  
会议地点:  
Singapore, SINGAPORE  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
2352-4847  
入藏号:  
WOS:001124237300038  
Indexed Date:  
2024-01-14  
Institution Address:  
North China Elect Power Univ, Sch Elect & Elect Engn, Baoding 071003, Peoples R China  
Institution Address:  
Sate Grid Sichuan Elect Power Res Inst, Chengdu 610000, Peoples R China  
Institution Address:  
Aostar Informat Technol Co Ltd, Chengdu 610074, Peoples R China  
Institution:  
Institution:  
Sate Grid Sichuan Elect Power Res Inst  
Institution:  
Aostar Informat Technol Co Ltd

### 414. Material structure segmentation method based on graph attention

摘要: With the development and integration of multiple disciplines, the integration of computer vision and materials science has greatly changed the original materials research methods. Existing methods can effectively segment the image of a specific scene, but there is no general method to segment and analyze the image of material accurately. To solve the problems of complex texture, blurred boundary and low contrast in material image, we propose a method that relies on multidimensional feature fusion to train the network more effectively with limited and available annotation samples. The architecture consists of an encoder, a graph attention module, a multi-scale feature fusion module and a decoder. We show that such a network can be trained end -to-end from the image. In electron microscope image, the segmentation results are superior to many previous advanced methods. Using this method, we can accurately identify multiple structures in material images, which provides important insights for multiphase segmentation of material images and searching for new mechanisms of structural transformation in material science.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
2352-4928  
入藏号:  
WOS:001040290600001  
Indexed Date:  
2023-08-15  
Institution Address:  
Shanghai Univ, Sch Comp Engn & Sci, 99 Shangda Rd, Shanghai 200444, Peoples R China  
Institution Address:  
Shanghai Univ, Key Lab Silicate Cultural Rel Conservat, Minist Educ, Shanghai, Peoples R China  
Institution Address:  
Zhejiang Lab, Hangzhou 311100, Peoples R China  
Institution:  
Institution:  
Institution:

### 415. Underwater Image Enhancement Method Based on Feature Fusion Neural

摘要: Aiming at the problems of uneven illumination of underwater image caused by supplementary illumination in deep-sea and night waters, image noise, low contrast and color deviation caused by suspended particles in water, a new underwater image enhancement method under non-uniform illumination is proposed. The heterogeneous feature fusion module is designed to fuse different levels and different levels of features, so as to improve the overall perception ability of the network to detail information and semantic information. Secondly, a new feature attention mechanism is designed to improve the traditional channel attention mechanism, and the improved channel attention and pixel attention mechanism are added to the heterogeneous feature fusion process to strengthen the ability of the network to extract pixel features with different turbidity. Then, the dynamic feature enhancement module is designed to adaptively expand the receptive field to improve the adaptability of the network to the image distortion scene and the ability of model conversion, and strengthen the network's learning of the region of interest. Finally, the color loss function is designed, and the absolute error loss and structural similarity loss are jointly minimized to correct the color deviation on the basis of maintaining the image texture. A multi-scale feature extraction module is designed to extract different levels of features at the beginning of the network, and the output results are obtained through the convolution layer with jump connection and the attention module. The experimental results on several data sets show that this method can have good results in processing synthetic underwater images and real underwater images, and can better restore the image color and texture details compared with the existing methods. It conforms to the characteristics of human vision, and the visual effect is better than the existing underwater image enhancement algorithms.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
2169-3536  
入藏号:  
WOS:000868713600001  
Indexed Date:  
2022-10-23  
Institution Address:  
Shandong Univ, Sch Mech Elect & Informat Engn, Weihai 264209, Shandong, Peoples R China  
Institution Address:  
59 Inst China Ordnance Ind, Chongqing 401331, Peoples R China  
Institution Address:  
Chongqing Res Inst Business Econ, Chongqing 401331, Peoples R China  
Institution:  
Institution:  
59 Inst China Ordnance Ind  
Institution:  
Chongqing Res Inst Business Econ

### 416. LEF-YOLO: a lightweight method for intelligent detection of four extreme

摘要: Background Extreme wildfires pose a serious threat to forest vegetation and human life because they spread more rapidly and are more intense than conventional wildfires. Detecting extreme wildfires is challenging due to their visual similarities to traditional fires, and existing models primarily detect the presence or absence of fires without focusing on distinguishing extreme wildfires and providing warnings.Aims To test a system for real time detection of four extreme wildfires.Methods We proposed a novel lightweight model, called LEF-YOLO, based on the YOLOv5 framework. To make the model lightweight, we introduce the bottleneck structure of MobileNetv3 and use depthwise separable convolution instead of conventional convolution. To improve the model's detection accuracy, we apply a multiscale feature fusion strategy and use a Coordinate Attention and Spatial Pyramid Pooling-Fast block to enhance feature extraction.Key results The LEF-YOLO model outperformed the comparison model on the extreme wildfire dataset we constructed, with our model having excellent performance of 2.7 GFLOPs, 61 FPS and 87.9% mAP.Conclusions The detection speed and accuracy of LEF-YOLO can be utilised for the real-time detection of four extreme wildfires in forest fire scenes.Implications The system can facilitate fire control decision-making and foster the intersection between fire science and computer science.  
We tested a lightweight architecture called LEF-YOLO for detecting four extreme wildfires. We found improved detection accuracy through multi-scale fusion and attention mechanism, and constructed four extreme wildfire datasets and compared these with multiple object detection models and lightweight feature extraction networks. This method is beneficial for the development of extreme wildfire field robots.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
5  
Web of Science 核心合集中的 "被引频次":  
51  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
51  
ISSN:  
1049-8001  
eISSN:  
1448-5516  
入藏号:  
WOS:001205875600006  
Indexed Date:  
2024-04-30  
Institution Address:  
Fuzhou Univ, Coll Phys & Informat Engn, Fuzhou 350116, Peoples R China  
Institution Address:  
Northeast Forestry Univ, Coll Mech & Elect Engn, Harbin, Peoples R China  
Institution Address:  
Fuzhou Univ, Zijin Sch Geol & Min, Fuzhou 350116, Peoples R China  
Institution Address:  
Huadong Engn Corp Ltd, Hangzhou 311122, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Huadong Engn Corp Ltd

### 417. FFCA-YOLO for Small Object Detection in Remote Sensing Images

摘要: Issues, such as insufficient feature representation and background confusion, make detection tasks for small object in remote sensing arduous. Particularly, when the algorithm will be deployed on board for real-time processing, which requires extensive optimization of accuracy and speed under limited computing resources. To tackle these problems, an efficient detector called feature enhancement, fusion and context aware YOLO (FFCA-YOLO) is proposed in this article. FFCA-YOLO includes three innovative lightweight and plug-and-play modules: feature enhancement module (FEM), feature fusion module (FFM), and spatial context aware module (SCAM). These three modules improve the network capabilities of local area awareness, multiscale feature fusion, and global association cross channels and space, respectively, while trying to avoid increasing complexity as possible. Thus, the weak feature representations of small objects are enhanced and the confusable backgrounds are suppressed. Two public remote sensing datasets (VEDAI and AI-TOD) for small object detection and one self-built dataset (USOD) are used to validate the effectiveness of FFCA-YOLO. The accuracy of FFCA-YOLO reaches 0.748, 0.617, and 0.909 (in terms of mAP50) that exceeds several benchmark models and the state-of-the-art methods. Meanwhile, the robustness of FFCA-YOLO is also validated under different simulated degradation conditions. Moreover, to further reduce computational resource consumption while ensuring efficiency, a lite version of FFCA-YOLO (L-FFCA-YOLO) is optimized by reconstructing the backbone and neck of FFCA-YOLO based on partial convolution (PConv). L-FFCA-YOLO has faster speed, smaller parameter scale, and lower computing power requirement but little accuracy loss compared with FFCA-YOLO.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
84  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
被引频次合计:  
84  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001178186400007  
Indexed Date:  
2024-04-18  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut NUAA, Coll Astronaut, Nanjing 211106, Peoples R China  
Institution Address:  
Acad Mil Sci, Natl Innovat Inst Def Technol, Beijing 100071, Peoples R China  
Institution:  
Institution:  
Acad Mil Sci

### 418. Lightweight ViT with Multiscale Feature Fusion for Driving Risk Rating

摘要: Addressing the issue of inadequate dynamic object detection accuracy in current road driving warning systems, this study proposes the RepBF-YOLOv8 detection algorithm aimed at efficient risk identification. The backbone network of YOLOv8n is replaced with the lightweight RepViT architecture, which is more suitable for visual tasks. This replacement simplifies the traditional structure, reduces the complexity of the backbone network, maximizes performance enhancement, and minimizes latency. Additionally, the FPN in the neck section is upgraded to Bi-FPN, which reduces nodes and span connections and incorporates rapid normalization to achieve fast multi-scale feature fusion. For risk grading, the algorithm infers distances and collision times, categorizing detected objects into high, medium, and low-risk levels, and uses different colors to warn the driver. Comparative experimental results show that the optimized algorithm improves Precision by 1.7%, Recall by 2.3%, mAP@0.5 by 1.53%, and mAP@0.5:0.95 by 2.91%. In road tests, the risk warning system achieves a frame detection rate ranging from a minimum of 38.4 fps to a maximum of 59.0 fps. The detection confidence for various objects remains above 0.71, reaching as high as 0.98. Specifically, the "Car" confidence ranges from 0.81 to 0.98, demonstrating the accuracy and robustness of vehicle risk detection.  
This study introduces the RepBF-YOLOv8 algorithm, enhancing dynamic object detection for autonomous driving. Employing the lightweight RepViT and advanced Bi-FPN, offering refined risk assessment through efficient multi-scale feature fusion and rapid risk categorization, significantly improving precision and recall while maintaining real-time performance. The study validates the model across various traffic scenarios, demonstrating robustness, and adaptability. image  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2513-0390  
入藏号:  
WOS:001290624400001  
Indexed Date:  
2024-08-18  
Institution Address:  
Yancheng Inst Technol, Sch Automot Engn, Yancheng 224051, Peoples R China  
Institution Address:  
Jiangsu Univ, Sch Automot & Traff Engn, Zhenjiang 212013, Peoples R China  
Institution Address:  
Chinese Acad Sci, Anhui Inst Opt & Fine Mech, Hefei 230031, Peoples R China  
Institution:  
Institution:  
Institution:

### 419. Deep Learning-Based Technique for Remote Sensing Image Enhancement Using

摘要: The present study proposes a novel deep-learning model for remote sensing image enhancement. It maintains image details while enhancing brightness in the feature extraction module. An improved hierarchical model named Global Spatial Attention Network (GSA-Net), based on U-Net for image enhancement, is proposed to improve the model's performance. To circumvent the issue of insufficient sample data, gamma correction is applied to create low-light images, which are then used as training examples. A loss function is constructed using the Structural Similarity (SSIM) and Peak Signal-to-Noise Ratio (PSNR) indices. The GSA-Net network and loss function are utilized to restore images obtained via low-light remote sensing. This proposed method was tested on the Northwestern Polytechnical University Very-High-Resolution 10 (NWPU VHR-10) dataset, and its overall superiority was demonstrated in comparison with other state-of-the-art algorithms using various objective assessment indicators, such as PSNR, SSIM, and Learned Perceptual Image Patch Similarity (LPIPS). Furthermore, in high-level visual tasks such as object detection, this novel method provides better remote sensing images with distinct details and higher contrast than the competing methods.  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
被引频次合计:  
11  
eISSN:  
1424-8220  
入藏号:  
WOS:001151016300001  
PubMed ID:  
38276366  
Indexed Date:  
2024-02-05  
Institution Address:  
Yangtze Univ, Sch Comp Sci, Jingzhou 434023, Peoples R China  
Institution:

### 420. GAN for Semantic Image Synthesis With Laplacian Pyramid and Multi-Scale

摘要: Most GAN-based methods utilize semantic layouts as input for generating realistic images. However, these layouts primarily consist of object contours and often lack detailed information, leading to suboptimal image quality in the generated outputs. To address this limitation, we propose a novel GAN architecture called LMCGAN designed specifically for synthesizing high-quality images. LMCGAN introduces a generator network structured around the laplacian pyramid, enabling the simultaneous generation of multi-scale feature maps.This approach allows the model to capture finer details at different resolutions, enhancing the overall realism of the generated images.To further improve the utilization of semantic maps, we integrate a multi-scale channel attention (MSCA) mechanism.This mechanism effectively focuses on channel-specific information in complex scenes, which is crucial for preserving essential details that may otherwise be lost. During the feature fusion phase, we implement a feature fusion block (FFBL) that is designed to capture important relationships across various scales. This block facilitates the integration of information from different resolutions, ensuring that the final output retains critical features. Additionally, we adopt a combination of conditional and unconditional methods to reduce noise during the training process, leading to more stable and effective training dynamics. Extensive experiments conducted on challenging datasets demonstrate that LMCGAN significantly outperforms existing methods in terms of both visual quality and quantitative evaluation metrics. The results indicate that our architecture not only generates more realistic images but also excels in preserving intricate details, marking a substantial advancement in the field of image synthesis using GANs.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001373836200014  
Indexed Date:  
2025-01-09  
Institution Address:  
Hubei Univ Technol, Sch Comp Sci, Wuhan 430068, Peoples R China  
Institution:

### 421. PerceptGuide: A Perception Driven Assistive Mobility Aid Based on

摘要: The paper introduces a novel wearable aid, PerceptGuide to help for visually impaired individuals to perceive scene around them. It is designed as a wearable, light weight chest rig bag, that incorporates a monocular camera, ultrasonic sensors, vibration motors, and a mono-earphone, powered by an embedded Nvidia Jetson development board. The system provides directional obstacle alerts through the vibration motors, allowing users to avoid obstacles on their path. A user-friendly push-button enables user to inquire about scene information in front of them. The scene details are effectively conveyed through a novel scene understanding approach, that combines multi-scale feature fusion, self-attention models, and a multilayer GRU (Gated Recurrent Unit) architecture on the ResNet50 backbone. The proposed system generates coherent and descriptive captions by capturing image features at different scales, enhancing the quality and contextual understanding of the scene details. The self-attention in both the encoder (ResNet50 + Feature fusion model) and decoder (multilayer GRU), effectively captures long-range dependencies and attend to relevant image regions. The quantitative evaluations conducted on the MSCOCO and Flicker8k datasets show the effectiveness of the model with improved Bleu-67.7, RougeL - 47.6, Meteor - 22.7 and CIEDR-67.4 scores. The PerceptGuide system exhibits exceptional real-time performance, generating audible captions in just 1.5 to 2 seconds. This rapid response time significantly aids visually impaired individuals in understanding the scenes around them. The qualitative evaluation of the aid emphasizes its real-time performance, demonstrating the generation of context-aware, semantically meaningful captions. This validates its potential as a wearable assistive aid for visually impaired people, with the added advantages of low power consumption, compactness, and a lightweight design.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001071719300001  
Indexed Date:  
2023-10-10  
Institution Address:  
Vishwakarma Inst Technol, Pune 411037, India  
Institution Address:  
Marathwada Mitra Mandals Coll Engn, Pune 411052, India  
Institution:  
Vishwakarma Inst Technol  
Institution:  
Marathwada Mitra Mandals Coll Engn

### 422. Research into the Applications of a Multi-Scale Feature Fusion Model in

摘要: Due to the increasing severity of aging populations in modern society, the accurate and timely identification of, and responses to, sudden abnormal behaviors of the elderly have become an urgent and important issue. In the current research on computer vision-based abnormal behavior recognition, most algorithms have shown poor generalization and recognition abilities in practical applications, as well as issues with recognizing single actions. To address these problems, an MSCS-DenseNet-LSTM model based on a multi-scale attention mechanism is proposed. This model integrates the MSCS (Multi-Scale Convolutional Structure) module into the initial convolutional layer of the DenseNet model to form a multi-scale convolution structure. It introduces the improved Inception X module into the Dense Block to form an Inception Dense structure, and gradually performs feature fusion through each Dense Block module. The CBAM attention mechanism module is added to the dual-layer LSTM to enhance the model's generalization ability while ensuring the accurate recognition of abnormal actions. Furthermore, to address the issue of single-action abnormal behavior datasets, the RGB image dataset RIDS (RGB image dataset) and the contour image dataset CIDS (contour image dataset) containing various abnormal behaviors were constructed. The experimental results validate that the proposed MSCS-DenseNet-LSTM model achieved an accuracy, sensitivity, and specificity of 98.80%, 98.75%, and 98.82% on the two datasets, and 98.30%, 98.28%, and 98.38%, respectively.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:001287220600001  
PubMed ID:  
39124111  
Indexed Date:  
2024-08-15  
Institution Address:  
Hebei Agr Univ, Sch Informat Sci & Technol, Baoding 071001, Peoples R China  
Institution Address:  
Hebei Key Lab Agr Big Data, Baoding 071001, Peoples R China  
Institution:  
Institution:  
Hebei Key Lab Agr Big Data

### 423. Hyperspectral image classification based on octave convolution and

摘要: Hyperspectral Image (HSI) classification is one of the main research directions of remote sensing applications. With the high dimension, strong correlation, and a large amount of HSI data, conventional classification methods often have problems like computation complexity, low classification accuracy, and poor generalization ability while extracting remote sensing image features. Moreover, the extraction of efficient features will impact the classification results. This work proposed an HSI classification method based on octave convolution and multi scale feature fusion. The low-level features of HSI were extracted by octave convolution, and the attention mechanism was introduced into the spatial and spectral dimensions to focus on the area of interest. Then, the spectral-spatial fusion features were extracted for the classification task. Based on 300 training samples, several experiments were carried out, and the results showed that the classification accuracy of the proposed method was 99.63%, 97.90%, and 98.69% on the widely used WHU-Hi dataset. The notable observation was that fusion of features of different scales helped improve the classification performance.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0141-6359  
eISSN:  
1873-2372  
入藏号:  
WOS:000806807400002  
Indexed Date:  
2022-06-15  
Institution Address:  
Sichuan Agr Univ, Sch Informat Engn, Yaan 625014, Peoples R China  
Institution Address:  
Sichuan Key Lab Agr Informat Engn, Yaan 625000, Sichuan, Peoples R China  
Institution:  
Institution:  
Sichuan Key Lab Agr Informat Engn

### 424. Attention-guided and resource-saving modules for semantic segmentation

摘要: Self-attention has been proven to be a quite powerful yet calculation-intensive method for scene semantic segmentation. Even though many efforts have been made to explore more effective and resource-saving ways to apply self-attention, there is still space in reducing the calculation consumption. Meanwhile, since self-attention is good at fusing information, its application should be extended to multi-scale-feature-fusion, which is barely researched while the information exchange paths between features in different resolutions are mostly addition and concatenation. A special partition method decreasing the computational complexity of self-attention is investigated, and a multi-scale-feature-attention (MFA) module fusing low-resolution features containing semantic information with high-resolution features having detailed information is presented at the same time. To be specific, the proposed multi-scale-partition-attention (MPA) module and MFA module are inserted into the backbone in sequence to fuse information among all the pixels in one highly extracted feature and the pixels from features with different resolutions, respectively. Extensive experiments are carried out on semantic segmentation benchmarks including PASCAL-Context and Cityscapes to demonstrate that these two improved modules can improve the performance of the backbone in scene semantic segmentation tasks that contain multiple classes and objects in both big and small sizes. (c) 2023 SPIE and IS&T  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:001075355800011  
Indexed Date:  
2024-03-09  
Institution Address:  
Harbin Inst Technol, Sch Mechatron Engn, Weihai, Peoples R China  
Institution Address:  
Army Acad Armored Forces, Vehicle Engn Dept, Beijing, Peoples R China  
Institution:  
Institution:  
Army Acad Armored Forces

### 425. Multi-scale feature fusion pyramid attention network for single image

摘要: Texture and color distortion are common in existing learning-based dehazing algorithms, and it is argued that one of the major reasons is that the shallow features of fog images are underutilized, and the deep features of fog images are insufficient for single image dehazing. In order to provide more texture and color information for image restoration, more shallow features need to be added in the process of image decoding. Therefore, a multi-scale feature fusion pyramid attention network (PAN) for single image dehazing is proposed. In PAN, combined with the attention mechanism, a shallow and deep feature fusion (SDF) strategy is designed. SDF considers multi-scale as well as channel-level fusion to provide feature information under different receptive fields while also highlighting important channels, such as texture and color information. DC is designed as a latent space mapping module to learn a mapping relationship between the latent space representation of the hazy image at low resolution and the corresponding latent space representation of the haze-free image. Additionally, network deconvolution (ND) and deformed convolution network (DCN) are introduced into PAN. The ND module can remove pixel-wise and channel-wise correlation of features, reduce data redundancy to obtain sparse representation of features, and speed up network convergence. The DCN module can use its adaptive receptive field to focus on the area of interest for calculation and play a role in texture feature enhancement. Finally, the perceptual loss is chosen as the regularization item of the loss function, which makes style features of the restored image closer to the real fog-free image. Extensive experiments reveal that the proposed PAN outperforms other existing dehazing methods on real-world and synthetic datasets.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1751-9659  
eISSN:  
1751-9667  
入藏号:  
WOS:000985022400001  
Indexed Date:  
2023-05-31  
Institution Address:  
Qufu Normal Univ, Sch Cyber Sci & Engn, Jining, Peoples R China  
Institution Address:  
Qufu Normal Univ, Sch Comp Sci, Jining, Peoples R China  
Institution Address:  
Qufu Normal Univ, Sch Comp Sci, Jining 276827, Peoples R China  
Institution:  
Institution:  
Institution:

### 426. RMANet: Refined-mixed attention network for progressive low-light image

摘要: Multi-scale feature fusion has been recognized as an effective strategy to boost the quality of low-light images. However, most existing methods directly extract multi-scale contextual information from severely degraded and down-sampled low-light images, resulting in a large amount of unexpected noise and degradation contaminating the learned multi-scale features. Moreover, there exist large redundant and overlapping features when directly concatenating multi-scale feature maps, which fails to consider different contributions of different scales. To conquer the above challenges, this paper presents a novel approach termed progressive Refined-Mixed Attention Network (RMANet) for low-light image enhancement. The proposed RMANet first targets a single-scale pre-enhancement and then progressively increases multi-scale spatial-channel attention fusion in a coarse-to-fine fashion. Additionally, we elaborately devise a Refined-Mixed Attention Module (RMAM) to first learn a parallel spatial-channel dominant features and then selectively integrate dominant features in the spatial and channel dimensions across multiple scales. Noticeably, our proposed RMANet is a lightweight yet flexible end-to-end framework that adapts to diverse application scenarios. Thorough experiments carried out upon three popular benchmark databases demonstrate that our approach surpasses existing methods in terms of both quantitative quality metrics and visual quality assessment.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0165-1684  
eISSN:  
1872-7557  
入藏号:  
WOS:001314263300001  
Indexed Date:  
2024-09-23  
Institution Address:  
Xian Polytech Univ, Sch Elect & Informat, Xian 710048, Peoples R China  
Institution Address:  
Xian Polytech Univ, Sch Comp Sci, Shaanxi Key Lab Clothing Intelligence & State Prov, Xian 710048, Peoples R China  
Institution Address:  
Xian Polytech Univ, Res Ctr Adv Networking & Intelligent Informat Serv, Sch Comp Sci, Xian 710048, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Chongqing Key Lab Image Cognit, Chongqing 400065, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 427. Real-Time Robotic Multigrasp Detection Using Anchor-Free Fully

摘要: Robotic grasping is essential for intelligent manufacturing. This article presents a novel anchor-free grasp detector based on fully convolutional network for detecting multiple valid grasps from RGB-D images in real time. Grasp detection is formulated as a closest horizontal or vertical rectangle regression task and a grasp angle classification task. By directing predicting grasps at feature points, our method eliminates the predefined anchors that commonly used in prior methods, and thus anchor-related hyperparameters and complex computations are avoided. For suppressing ambiguous and low-quality training samples, a new sample assignment strategy that combines center-sampling and regression weights is proposed. Our method achieves a state-of-the-art accuracy of 99.4% on Cornell and 96.2% on Jacquard dataset, and real-time speed of 104 frames per second, with approximately 2 x fewer parameters and 8x less training time compared to previous one-stage detector. Moreover, an efficient multi-scale feature fusion module is integrated to improve the performance of multigrasp detection by 25%. In real-world robotic grasping of novel objects, our method achieves a grasp success rate of 91.3% for single object and 83.3% for multiple objects with only 26 ms used for the whole planning. The results demonstrate that our method is robust for potential industrial applications.  
Web of Science 核心合集中的 "被引频次":  
30  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
31  
ISSN:  
0278-0046  
eISSN:  
1557-9948  
入藏号:  
WOS:000838702800096  
Indexed Date:  
2022-09-28  
Institution Address:  
Harbin Inst Technol, State Key Lab Robot & Syst, Harbin 150001, Peoples R China  
Institution:

### 428. Regression-Guided Refocusing Learning With Feature Alignment for Remote

摘要: Tiny object detection is a formidable challenge in remote sensing intelligent interpretation. Tiny objects are usually fuzzy, densely distributed, and highly sensitive to positioning errors, which leads to the mainstream detector usually achieving suboptimal detection performance when facing tiny objects. To address the mismatch of mainstream detector architectures and model optimization strategies in the context of tiny object detection, this article presents an efficient and interpretable algorithm for tiny object detection, termed the cross-attention-based feature fusion enhanced tiny object detection network (CAF(2)ENet). First, the cross-attention mechanism is introduced to refine the upsampling results of deep features. This refinement improves the precision of multiscale feature fusion. Second, a training strategy named regression-based refocusing learning is introduced. Deviating from the conventional optimization strategy, our method guides the optimizer to prioritize higher-quality detection boxes by adjusting sample weights. This adjustment significantly amplifies the detector's potential to achieve superior detection results. Finally, the object composite confidence score is employed for the interpretable filtering of detection boxes. Extensive experiments on tiny object detection in aerial images (AI-TOD) and object detection in optical remote sensing images (DIOR) datasets are carried out, and comparison indicates that the proposed CAF(2)ENet can perform the remarkable performance compared to other state-of-the-art (SOTA) tiny object detection detectors, as it can reach 63.7% average precision (AP(50)) on AI-TOD and 75.4% AP(50 )on DIOR, achieve SOTA performance.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001246585400001  
Indexed Date:  
2024-06-27  
Institution Address:  
Beijing Inst Technol, Natl Key Lab Sci & Technol Space Born Intelligent, Beijing 100081, Peoples R China  
Institution Address:  
Peking Univ, Sch Comp Sci, Beijing 100871, Peoples R China  
Institution Address:  
Peking Univ, Ctr Frontiers Comp Studies CFCS, Beijing 100871, Peoples R China  
Institution:  
Institution:  
Institution:

### 429. Underwater image enhancement via multiscale disentanglement strategy

摘要: Underwater images suffer from color casts, low illumination, and blurred details caused by light absorption and scattering in water. Existing data-driven methods often overlook the scene characteristics of underwater imaging, limiting their expressive power. To address the above issues, we propose a Multiscale Disentanglement Network (MD-Net) for Underwater Image Enhancement (UIE), which mainly consists of scene radiance disentanglement (SRD) and transmission map disentanglement (TMD) modules. Specifically, MD-Net first disentangles original images into three physical parameters which are scene radiance (clear image), transmission map, and global background light. The proposed network then reconstructs these physical parameters into underwater images. Furthermore, MD-Net introduces class adversarial learning between the original and reconstructed images to supervise the disentanglement accuracy of the network. Moreover, we design a multi-level fusion module (MFM) and dual-layer weight estimation unit (DWEU) for color cast adjustment and visibility enhancement. Finally, we conduct extensive qualitative and quantitative experiments on three benchmark datasets, which demonstrate that our approach outperforms other traditional and state-of-the-art methods. Our code and results are available at: https://github.com/WYJGR/MD-Net.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2045-2322  
入藏号:  
WOS:001426710900032  
PubMed ID:  
39971756  
Indexed Date:  
2025-02-26  
Institution Address:  
Minjiang Univ, Sch Comp & Big Data, Fujian Prov Key Lab Informat Proc & Intelligent Co, Fuzhou 350121, Peoples R China  
Institution Address:  
Wenzhou Med Univ, Sch Ophthalmol & Optometry, Wenzhou 325000, Peoples R China  
Institution Address:  
Univ Lahore, Dept Comp Engn, Lahore 54000, Punjab, Pakistan  
Institution Address:  
Hanshan Normal Univ, Sch Phys & Elect Engn, Chaozhou 521041, Guangdong, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 430. MarsSeg: Mars Surface Semantic Segmentation With Multilevel Extractor

摘要: The segmentation and interpretation of the Martian surface play a pivotal role in Mars exploration, providing essential data for the trajectory planning and obstacle avoidance of rovers. However, the complex topography, self-similar surface features, and the lack of extensive annotated data pose significant challenges to the high-precision semantic segmentation of the Martian surface. To address these challenges, we propose a novel encoder-decoder-based Mars segmentation network, termed MarsSeg. To facilitate a high-level semantic understanding across the multilevel feature maps, we introduce a feature enhancement module, which incorporates a multiscale feature pyramid (MFP) and strip attention pyramid pooling module (SAPPM). The MFP is specifically designed for shallow feature enhancement, thereby enabling the expression of local details and small objects. Conversely, the SAPPM is employed for deep feature enhancement, facilitating the extraction of high-level semantic category-related information. To effectively fuse features from different levels, we propose a feature fusion module, which contains Mars polarized self-attention (Mars-PSA) and pixel attention head (PA-Head). Mars-PSA enables the fusion of multilevel information while directing the model's attention to salient features. The PA-Head focuses on detailed information at the pixel level. Experimental results derived from the Mars-Seg and AI4Mars datasets prove that the proposed MarsSeg outperforms other state-of-the-art methods in segmentation performance, validating the efficacy of each proposed component.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001410640900011  
Indexed Date:  
2025-02-14  
Institution Address:  
Beihang Univ, Image Proc Ctr, Sch Astronaut, Beijing 100191, Peoples R China  
Institution Address:  
Beihang Univ, State Key Lab Virtual Real Technol & Syst, Beijing 100191, Peoples R China  
Institution Address:  
Shanghai Artificial Intelligence Lab, Shanghai 200232, Peoples R China  
Institution:  
Institution:  
Institution:  
Shanghai Artificial Intelligence Lab

### 431. A Multi-Scale Liver Tumor Segmentation Method Based on Residual and

摘要: Liver cancer is one of the malignancies with high mortality rates worldwide, and its timely detection and accurate diagnosis are crucial for improving patient prognosis. To address the limitations of traditional image segmentation techniques and the U-Net network in capturing fine image features, this study proposes an improved model based on the U-Net architecture, named RHEU-Net. By replacing traditional convolution modules in the encoder and decoder with improved residual modules, the network's feature extraction capabilities and gradient stability are enhanced. A Hybrid Gated Attention (HGA) module is integrated before the skip connections, enabling the parallel processing of channel and spatial attentions, optimizing the feature fusion strategy, and effectively replenishing image details. A Multi-Scale Feature Enhancement (MSFE) layer is introduced at the bottleneck, utilizing multi-scale feature extraction technology to further enhance the expression of receptive fields and contextual information, improving the overall feature representation effect. Testing on the LiTS2017 dataset demonstrated that RHEU-Net achieved Dice scores of 95.72% for liver segmentation and 70.19% for tumor segmentation. These results validate the effectiveness of RHEU-Net and underscore its potential for clinical application.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001311394000001  
PubMed ID:  
39275756  
Indexed Date:  
2024-09-18  
Institution Address:  
Changchun Univ, Coll Comp Sci & Technol, 6543 Satellite Rd, Changchun 130022, Peoples R China  
Institution:

### 432. LFF-YOLO: A YOLO Algorithm With Lightweight Feature Fusion Network for

摘要: The detection of defects is indispensable in industrial production. Surface defects have different scales. Both minimal flaws and significant scratches may appear on the same product. The standard method uses a multi-scale feature fusion network, introducing many parameters that may reduce the inference speed. In actual industrial production scenarios, inference speed and accuracy play an equally important role. Therefore we propose an algorithm to effectively improve the detection speed while improving the detection accuracy. The model proposed in this paper called "YOLO with lightweight feature fusion network (LFF-YOLO). " First, we use ShuffleNetv2 as a feature extraction network to reduce the number of parameters. Then, to improve the efficiency of multi-scale feature fusion, we propose the lightweight feature pyramid network (LFPN). Considering that the fixed receptive field is difficult to adapt to the defects of different scales, it may lead to the difficulty of model convergence and seriously affect the detection performance. Therefore, we propose the adaptive receptive field feature extraction (ARFFE) module, which weights the multi-receptive field channels to generate multi-receptive field information. In addition, focal loss is used to solve the problem of imbalance between positive and negative samples. Finally, we conducted experiments on NEU-DET (79.23% mAP), Peking University printed circuit board defect dataset (93.31% mAP),and GC10-DET (59.78% mAP), respectively. Extensive experiments show that our proposed method achieves optimal detection speed compared with the prevailing methods, and the detection accuracy of our method is also highly competitive. We open-soure our code in the following URL:https://github.com/syyang2022/LFF-YOLO  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
41  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
44  
ISSN:  
2169-3536  
入藏号:  
WOS:000902044800001  
Indexed Date:  
2023-01-10  
Institution Address:  
Zhejiang Univ Sci & Technol, Dept Elect Informat Engn, Hangzhou 310023, Zhejiang, Peoples R China  
Institution Address:  
Zhejiang Dingli Ind Co Ltd, Lishui 321400, Zhejiang, Peoples R China  
Institution:  
Institution:  
Zhejiang Dingli Ind Co Ltd

### 433. Multiscale Neighborhood Cluster Scene Flow Prior for LiDAR Point Clouds

摘要: Scene flow estimation, which aims to predict point-wise displacement in 3-D space from sequential data, is a challenging task with wide application in fields such as robotics and autonomous driving. Currently, the accuracy of scene flow estimation from sparse point clouds using prior-based models is suboptimal. Therefore, we revisit the point-by-point scene flow prior and propose a multiscale neighborhood cluster scene flow prior (MNCSFP) to enhance the accuracy of scene flow estimation in sparse point clouds. We optimize the prior model utilizing the multiscale neighborhood cluster feature of the point cloud, in which the point neighborhood is constructed only once. According to the neighborhood index (NI), we design a multiscale neighborhood cluster feature construction (MNCFC) module. The MNCFC module rapidly constructs multiscale neighborhood cluster features using a Gaussian-based neighborhood feature normalization (GNFN) strategy to improve the representation of neighborhood cluster characteristics. Moreover, we propose a neighborhood cluster weighted aggregation (NCWA) module to encode neighborhood cluster features. In NCWA, we design the logsoft function to calculate the neighborhood cluster weights and complete the extraction and aggregation of neighborhood cluster features. Furthermore, we design the multiscale feature fusion (MFF) module that combines the symmetry operation and the logsoft function to complete the fusion of multiscale features to enhance the sparse point feature stability. We evaluate our method on samples from the KITTI, Argoverse, nuScenes, and Waymo Open datasets and demonstrate that it outperforms existing methods and achieves advanced performance.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001397823700031  
Indexed Date:  
2025-01-22  
Institution Address:  
China Univ Min & Technol Beijing, Sch Artificial Intelligence, Beijing 100083, Peoples R China  
Institution:

### 434. A deep learning lightweight model for real-time captive macaque facial

摘要: Automated behavior monitoring of macaques offers transformative potential for advancing biomedical research and animal welfare. However, reliably identifying individual macaques in group environments remains a significant challenge. This study introduces ACE-YOLOX, a lightweight facial recognition model tailored for captive macaques. ACE-YOLOX incorporates Efficient Channel Attention (ECA), Complete Intersection over Union loss (CIoU), and Adaptive Spatial Feature Fusion (ASFF) into the YOLOX framework, enhancing prediction accuracy while reducing computational complexity. These integrated approaches enable effective multiscale feature extraction. Using a dataset comprising 179 400 labeled facial images from 1 196 macaques, ACE-YOLOX surpassed the performance of classical object detection models, demonstrating superior accuracy and real-time processing capabilities. An Android application was also developed to deploy ACE-YOLOX on smartphones, enabling on-device, real-time macaque recognition. Our experimental results highlight the potential of ACE-YOLOX as a non-invasive identification tool, offering an important foundation for future studies in macaque facial expression recognition, cognitive psychology, and social behavior.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2095-8137  
入藏号:  
WOS:001450479700008  
PubMed ID:  
40049662  
Indexed Date:  
2025-03-29  
Institution Address:  
Chinese Acad Sci, Kunming Inst Zool, Key Lab Genet Evolut & Anim Models, Kunming Nat Hist Museum Zool, Kunming 650201, Yunnan, Peoples R China  
Institution Address:  
Yunnan Agr Univ, Coll Big Data, Kunming 650201, Yunnan, Peoples R China  
Institution Address:  
Chinese Acad Sci, Natl Resource Ctr Nonhuman Primates, Kunming Primate Res Ctr, Kunming 650107, Yunnan, Peoples R China  
Institution Address:  
Chinese Acad Sci, Kunming Inst Zool, Natl Res Facil Phenotyp & Genet Anal Model Anim, Primate Facil, Kunming 650107, Yunnan, Peoples R China  
Institution Address:  
Yunnan Key Lab Biodivers Informat, Kunming 650223, Yunnan, Peoples R China  
Institution Address:  
Chinese Acad Sci, Ctr Excellence Anim Evolut & Genet, Kunming 650201, Yunnan, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Yunnan Key Lab Biodivers Informat  
Institution:

### 435. ELLK-Net: An Efficient Lightweight Large Kernel Network for SAR Ship

摘要: ELLK-Net, an efficient, lightweight network with a large kernel, is proposed for synthetic aperture radar (SAR) ship detection. It addresses background variations, different ship scales, and noise interference challenges. ELLK-Net uses an anchor-free detector framework and sequentially decomposes large kernel convolutions to capture comprehensive global information and long-range dependencies. It adaptively selects convolution kernels on the basis of target characteristics, enhancing multiscale feature expression. A novel large kernel multiscale attention (LKMA) module is introduced to enhance interlayer feature fusion and semantic alignment, mitigating the impacts of overlapping ships and scattering noise. Structural reparameterization techniques optimize inference speed across devices without compromising accuracy. The experimental results on the SAR ship detection dataset (SSDD) and high-resolution SAR image dataset (HRSID) datasets demonstrate that ELLK-Net achieves impressive AP50 values of 95.6% and 90.6% for horizontal box detection and 89.7% and 79.7% for rotating box detection, respectively. The reparameterized detector exhibits a significant 48.7% FPS improvement on the Nvidia Jetson NX platform, indicating its suitability for edge computing deployment. The code is available at https://github.com/CHD-IPAC/ELLK-Net.  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001315975400015  
Indexed Date:  
2024-09-25  
Institution Address:  
Changan Univ, Sch Elect & Control Engn, Xian 710064, Peoples R China  
Institution:

### 436. Multiscale Feature Fusion for the Multistage Denoising of Airborne

摘要: Compared with the existing modes of LiDAR, single-photon LiDAR (SPL) can acquire terrain data more efficiently. However, influenced by the photon-sensitive detectors, the collected point cloud data contain a large number of noisy points. Most of the existing denoising techniques are based on the sparsity assumption of point cloud noise, which does not hold for SPL point clouds, so the existing denoising methods cannot effectively remove the noisy points from SPL point clouds. To solve the above problems, we proposed a novel multistage denoising strategy with fused multiscale features. The multiscale features were fused to enrich contextual information of the point cloud at different scales. In addition, we utilized multistage denoising to solve the problem that a single-round denoising could not effectively remove enough noise points in some areas. Interestingly, the multiscale features also prevent an increase in false-alarm ratio during multistage denoising. The experimental results indicate that the proposed denoising approach achieved 97.58%, 99.59%, 95.70%, and 77.92% F1-scores in the urban, suburban, mountain, and water areas, respectively, and it outperformed the existing denoising methods such as Statistical Outlier Removal. The proposed approach significantly improved the denoising precision of airborne point clouds from single-photon LiDAR, especially in water areas and dense urban areas.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
eISSN:  
2072-4292  
入藏号:  
WOS:000908710500001  
Indexed Date:  
2023-01-22  
Institution Address:  
Southwest Jiaotong Univ, Fac Geosci & Environm Engn, Chengdu 611756, Peoples R China  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying, Mapping & Remote Sensing, Wuhan 430072, Peoples R China  
Institution Address:  
China Acad Railway Sci Grp Co Ltd, Beijing 100081, Peoples R China  
Institution Address:  
Beijing Jingwei Informat Technol Co Ltd, Beijing 100081, Peoples R China  
Institution:  
Institution:  
Institution:  
China Acad Railway Sci Grp Co Ltd  
Institution:  
Beijing Jingwei Informat Technol Co Ltd

### 437. Lightweight Traffic Classification Model Based on Deep Learning

摘要: The development of mobile computing and the Internet of Things (IoT) has led to a surge in traffic volume, which creates a heavy burden for efficient network management. The network management requires high computational overheads to make traffic classification, which is even worse when in edge networks; existing approaches sacrifice the efficiency to obtain high-precision classification results, which are no longer suitable for limited resources edge network scenario. Given the problem, existing traffic classification generally has huge parameters and especially computational complexity. We propose a lightweight traffic classification model based on the Mobilenetv3 and improve it for an ingenious balance between performance and lightweight. Firstly, we adjust the model scale, width, and resolution to substantially reduce the number of model parameters and computations. Secondly, we embed precise spatial information on the attention mechanism to enhance the traffic flow-level feature extraction capability. Thirdly, we use the lightweight multiscale feature fusion to obtain the multiscale flow-level features of traffic. Experiments show that our model has excellent classification accuracy and operational efficiency. The accuracy of the traffic classification model designed in our work has reached more than 99.82%, and the parameter and computation amount are significantly reduced to 0.26 M and 5.26 M. In addition, the simulation experiments on Raspberry Pi prove the proposed model can realize real-time classification capability in the edge network.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1530-8669  
eISSN:  
1530-8677  
入藏号:  
WOS:000876432500002  
Indexed Date:  
2022-11-13  
Institution Address:  
PLA Strateg Support Force Informat Engn Univ, Informat Technol Inst, Zhengzhou 450000, Peoples R China  
Institution Address:  
Purple Mt Lab, Endogenous Safety & Secur Res Ctr, Nanjing 211100, Peoples R China  
Institution:  
Institution:  
Purple Mt Lab

### 438. SalDA: DeepConvNet Greets Attention for Visual Saliency Prediction

摘要: Predicting salient regions in images requires the capture of contextual information in the scene. Conventional saliency models typically use the encoder-decoder architecture and multiscale feature fusion for modeling contextual features, which, however, possess huge computational cost and model parameters. In this article, we address the saliency prediction task by capturing long-range dependencies based on the self-attention mechanism. Self-attention has been widely used in image recognition or other classification tasks, but is still rarely being considered in regression-based saliency prediction task. Inspired by the nonlocal block, we propose a new saliency prediction network in which deep convolutional network is integrated with the attention mechanism, namely, SalDA. Considering each feature map may capture different salient regions, our spatial attention module first adaptively aggregates the feature at each position by a weighted sum of the features at all positions within each independent channel. Meanwhile, in order to capture interdependence between channels, we also introduce a channel attention module to integrate different features among different channels. We combine these two attention modules into a multiattention module to further improve the saliency map prediction for the network. We show the effectiveness of SalDA on the largest saliency prediction data set SALICON. Compared to other state-of-the-art methods in this area, we can yield comparable saliency prediction performance, but with substantially less model parameters and shorter inference time.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
2379-8920  
eISSN:  
2379-8939  
入藏号:  
WOS:001167556100031  
Indexed Date:  
2024-03-21  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Comp Sci & Technol, Nanjing 211106, Peoples R China  
Institution Address:  
ByteDance, Shanghai 201100, Peoples R China  
Institution:  
Institution:  
ByteDance

### 439. Small target detection algorithm for printing defects detection based on

摘要: Small target detection is an important research direction in the field of computer vision, which is widely used in popular fields such as industrial defect detection, satellite remote sensing image detection. However, in printing defects detection, due to the complex defect background and small target, it is difficult to extract multi-scale features, and the extracted features have less available information. Therefore, this paper improves YOLOv7 and proposes an end-to-end printing defects detection algorithm based on context structure perception and multi-scale feature fusion (CM-YOLOv7). CM-YOLOv7 is mainly composed of a Context Structure Awareness Module (CSAM), a Multi-scale Feature Interaction Module (MFIM) and a Feature Refinement Layer Module (FRLM). Firstly, the CSAM uses multiple convolution kernels of different sizes to obtain features of different receptive fields and enhance multi-scale feature extraction. Secondly, for the extracted multi-scale features, the MFIM adaptively fuses the features of adjacent layers to achieve mutual learning between coarse-grained information and fine-grained information and improves the expression ability of small target features lost after convolution. Finally, in order to refine the edge information of the defect target more effectively and enhance the feature expression ability of the image, an FRLM is designed in MFIM. In the experimental part, this paper utilizes the printing defects detection dataset and DOTA-V1.0 dataset in order to better evaluate the effect of the algorithm. The experimental results show that the CM-YOLOv7 model proposed in this paper has a more accurate detection effect on the small target of printing defects.  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001082621600001  
Indexed Date:  
2023-11-08  
Institution Address:  
Xian Univ Technol, Coll Fac Printing, Packaging Engn & Digital Media Technol, Xian 710048, Peoples R China  
Institution Address:  
Key Lab Printing & Packaging Engn Shaanxi Prov, Xian 710048, Peoples R China  
Institution:  
Institution:  
Key Lab Printing & Packaging Engn Shaanxi Prov

### 440. Multiscale Pillars Fusion for 4-D Radar Object Detection With Radar Data

摘要: The 4-D millimeter-wave (mmWave) radar has garnered significant interest in autonomous driving due to its robust ability to measure the range, azimuth, elevation, and velocity of targets. In addition, this is also attributed to its resilience in extreme environments, as well as lower costs. However, the sparsity and noise inherent in 4-D radar point clouds limit its perception of targets in practical application. To address this, we propose a radar point cloud enhancement (RPCE) method that leverages semantic knowledge from RGB images to enhance the quality of 4-D radar point cloud by generating virtual points and improving feature quality. Furthermore, we introduce MSPFNet, a multiscale pillars fusion network framework built on the enhanced point clouds. MSPFNet divides the point clouds into multiscale pillars and then extracts multiscale features from each pillar's points. For these multiscale features, we propose a multiscale feature fusion (MSFF) module that can simultaneously capture intra- and interrelationship between different pillars through self-attention and cross-attention manners. Finally, we quantitatively and qualitatively evaluate our approach using the View-of-Delft (VoD) dataset and conduct extensive ablation studies to investigate its effectiveness. Our method's performance is improved by 8.9% of mean average precision (mAP) in the entire annotated area and 5% of mAP in the driving corridor compared with the baseline. The experimental results and analysis indicate that the proposed method can effectively improve the performance for 4-D radar object detection with comparable inference time.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001414892100020  
Indexed Date:  
2025-02-14  
Institution Address:  
Suzhou Key Lab Microwave Imaging Proc & Applicat T, Suzhou 215123, Peoples R China  
Institution Address:  
Natl Key Lab Microwave Imaging, Beijing 100190, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 101408, Peoples R China  
Institution Address:  
Suzhou Aerosp Informat Res Inst, Suzhou 215123, Peoples R China  
Institution Address:  
Beijing Autord Technol Co Ltd, Beijing 100102, Peoples R China  
Institution Address:  
Natl Def Univ, Sch Natl Secur, Beijing 100091, Peoples R China  
Institution Address:  
Nanjing Vocat Inst Railway Technol, Nanjing 210008, Peoples R China  
Institution:  
Suzhou Key Lab Microwave Imaging Proc & Applicat T  
Institution:  
Natl Key Lab Microwave Imaging  
Institution:  
Institution:  
Institution:  
Suzhou Aerosp Informat Res Inst  
Institution:  
Beijing Autord Technol Co Ltd  
Institution:  
Natl Def Univ  
Institution:  
Nanjing Vocat Inst Railway Technol

### 441. Multiscale object detection in high-resolution remote sensing images via

摘要: Due to the complex background and spatial distribution, it brings great challenge to object detection in high-resolution remote sensing images. In view of the characteristics of various scales, arbitrary orientations, shape variations, and dense arrangement, a multiscale object detection method in high-resolution remote sensing images is proposed by using rotation invariance deep features driven by channel attention. First, a channel attention module is added to our feature fusion and scaling-based single shot detector (FS-SSD) to strengthen the long-term semantic dependence between objects for improving the discriminative ability of the deep features. Then, an oriented response convolution is followed to generate feature maps with orientation channels to produce rotation invariant deep features. Finally, multiscale objects are predicted in a high-resolution remote sensing image by fusing various scale feature maps with multiscale feature module in FS-SSD. Five experiments are conducted on NWPU VHR-10 dataset and achieve better detection performance compared with the state-of-the-art methods.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
19  
被引频次合计:  
21  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:000662247000001  
Indexed Date:  
2021-07-01  
Institution Address:  
Beijing Univ Technol, Fac Informat Technol, Beijing, Peoples R China  
Institution Address:  
Beijing Univ Technol, Beijing Key Lab Computat Intelligence & Intellige, Fac Informat Technol, Beijing, Peoples R China  
Institution Address:  
China Univ Geosci, Inst Math Geol Remote Sensing, Sch Earth Resources, Wuhan, Peoples R China  
Institution:  
Institution:  
Institution:

### 442. Remote Sensing Small Object Detection Network Based on Attention

摘要: In remote sensing images, small objects have too few discriminative features, are easily confused with background information, and are difficult to locate, leading to a degradation in detection accuracy when using general object detection networks for aerial images. To solve the above problems, we propose a remote sensing small object detection network based on the attention mechanism and multi-scale feature fusion, and name it AMMFN. Firstly, a detection head enhancement module (DHEM) was designed to strengthen the characterization of small object features through a combination of multi-scale feature fusion and attention mechanisms. Secondly, an attention mechanism based channel cascade (AMCC) module was designed to reduce the redundant information in the feature layer and protect small objects from information loss during feature fusion. Then, the Normalized Wasserstein Distance (NWD) was introduced and combined with Generalized Intersection over Union (GIoU) as the location regression loss function to improve the optimization weight of the model for small objects and the accuracy of the regression boxes. Finally, an object detection layer was added to improve the object feature extraction ability at different scales. Experimental results from the Unmanned Aerial Vehicles (UAV) dataset VisDrone2021 and the homemade dataset show that the AMMFN improves the APs values by 2.4% and 3.2%, respectively, compared with YOLOv5s, which represents an effective improvement in the detection accuracy of small objects.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
19  
被引频次合计:  
19  
eISSN:  
2072-4292  
入藏号:  
WOS:001004173300001  
Indexed Date:  
2023-06-21  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Automat, Xian Key Lab Adv Control & Intelligent Proc, Xian Robert Intelligent Syst Int Sci & Technol Coo, Xian 710121, Peoples R China  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Commun & Informat Engn, Xian 710121, Peoples R China  
Institution:  
Institution:

### 443. Parallel feature enhancement and adaptive weighted feature fusion for

摘要: To tackle the challenges posed by the insensitivity of current multi-scale networks to image detailed information and their limited capacity to model contextual relationships, our paper proposes a novel semantic segmentation network called LEFNet. It is based on Parallel Feature Enhancement (PFE) and Adaptive Weighted Feature Fusion (AWFF). In the coding stage, the PFE module performs posterior enhancement of multi-scale features by Detail Sharpening Attention (DSA) and High-level Dilation Fusion (HDF) methods. The DSA guides the learning of detailed information in low-level features. The HDF broadens the perceptual field, enabling the acquisition of richer high-level features. In the decoding stage, the AWFF module supersedes the conventional feature fusion methods. The AWFF module constructs perceptual factors for each multi-scale feature map, enabling weighted learning of features. It emphasizes features with stronger semantic information, making them more decisive in pixel classification. It integrates features more reasonably based on the relevance of global contextual information, fully releasing the expressive potential of encoded features. Our method achieves mIoU scores of 82.8%, 49.3%, and 45.4% on the Cityscapes, ADE 20 K, and COCO-Stuff 164 K datasets, respectively, reaching an advanced level on popular benchmarks. The experimental results show that LEFNet alleviates the challenge of insensitivity of multi-scale networks to image detailed information, improves its ability to model contextual relationships, and significantly improves segmentation performance.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1051-2004  
eISSN:  
1095-4333  
入藏号:  
WOS:001311469000001  
Indexed Date:  
2024-09-25  
Institution Address:  
Henan Univ, Sch Artificial Intelligence, Zhengzhou 450046, Peoples R China  
Institution:

### 444. Construction of multi-scale feature fusion segmentation model of MRI

摘要: BACKGROUND: Early diagnosis of knee osteoarthritis is an important area of research in the field of clinical medicine. Due to the complexity in the MRI imaging sequences and the diverse structure of cartilage, there are many challenges in the segmentation of knee bone and cartilage. Relevant studies have conducted semantic fusion processing through splicing or summing forms, which results in reduced resolution and the accumulation of redundant information.  
OBJECTIVE: This study was envisaged to construct an MRI image segmentation model to improve the diagnostic efficiency and accuracy of different grade knee osteoarthritis by adopting the Dual Attention and Multi-scale Feature Fusion Segmentation network (DA-MFFSnet).  
METHODS: The feature information of different scales was fused through the Multi-scale Attention Downsample module to extract more accurate feature information, and the Global Attention Upsample module weighted lower-level feature information to reduce the loss of key information.  
RESULTS: The collected MRI knee images were screened and labeled, and the study results showed that the segmentation effect of DA-MFFSNet model was closer to that of the manually labeled images. The mean intersection over union, the dice similarity coefficient and the volumetric overlap error was 92.74%, 91.08% and 7.44%, respectively, and the accuracy of the differential diagnosis of knee osteoarthritis was 84.42%.  
CONCLUSIONS: The model exhibited better stability and classification effect. Our results indicated that the Dual Attention and Multi-scale Feature Fusion Segmentation model can improve the segmentation effect of MRI knee images in mild and medium knee osteoarthritis, thereby offering an important clinical value and improving the accuracy of the clinical diagnosis.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0928-7329  
eISSN:  
1878-7401  
入藏号:  
WOS:001255718100024  
PubMed ID:  
38759056  
Indexed Date:  
2024-08-09  
Institution Address:  
Beijing Hlth Vocat Coll, Dept Med Tech, Beijing 102402, Peoples R China  
Institution Address:  
Capital Med Univ, Fuxing Hosp, Dept Radiol, Beijing, Peoples R China  
Institution:  
Beijing Hlth Vocat Coll  
Institution:

### 445. Semantic Segmentation of Remote Sensing Image Based on Multi-Scale

摘要: This paper analyzes the existed processing scheme, and proposes a multi-scale semantic encoder-decoder networks (MSEDNet) by comprehensively using multiple technologies for the problems in remote sensing image semantic segmentation both multi-level information extraction and multi-scale feature diagram dependence characteristic. The MSEDNet consists of two parts: encoding part and decoding part. In the encoding part, the enhanced MobileNetV3 with residuals coordinate spatial attention (RCSA) is firstly proposed to extract semantic information, and then a multi-layer enhanced semantic context module (ESCM) is designed to improve representation ability of the multi-scale structure feature map. In the decoding part, a strengthen spatial detail information module (SSDIM) based on Multi-core Convolution and Focus Parallel is proposed to enhance the details and structural information of shallow features. Then triplet iterative multiscale feature fusion (TIMSFF) strategy is designed to strengthen the multi-scale context fusion both deep global semantic information and shallow local detail features, for improving the segmentation accuracy. The proposed model has been experimentally verified on the ISPRS Vaihingen and Potsdam dataset. The overall segmentation accuracy (OA) reached 95.699% and 95.534% respectively, the mean F1-score (mF1) increased by 2.661% and 2.929% respectively, and the mean intersection over union (mIoU) increased by 3.973%and 4.012%, respectively. The number of param dropped to 6.77 M.  
摘要:  
针对遥感图像语义分割中存在的多层次信息提取和多尺度特征图上下文依赖性两个问题,本文分析现有处理方案,提出了一种综合运用多项技术的多尺度语义编解码网络(Multi-scale Semantic Encoder-Decoder Networks, MSEDNet). MSEDNet由编码与解码两部分构成.编码阶段,首先提出残差协同空间注意(Residuals Coordinate Spatial Attention,RCSA)的MobileNetV3增强型模块,提取语义信息;其次,设计多层增强语义上下文模块(Enhance Semantic Context Module,ESCM),提升多尺度结构特征图的表征能力.解码阶段,首先提出多核卷积与Focus并行的强化空间细节信息模块(Strengthen Spatial Detail Information Module,SSDIM),增强浅层特征细节和结构信息;其次,设计了三元迭代多尺度特征融合(Triplet Iterative Multi-Scale Feature Fusion,TIMSFF)策略,强化图像深层全局语义信息与浅层局部细节特征的多尺度融合,提升分割精度.所提模型在ISPRS Vaihingen和Potsdam数据集上验证,总体分割精度(Overall Accuracy,OA)分别达到95.699%、95.534%,平均F1-score(mean F1-score,mF1)分别提高2.661%和2.929%,且平均交并比(mean Intersection over Union,mIoU)分别增长3.973%和4.012%.所耗参数量Param下降至6.77 M.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0372-2112  
入藏号:  
CSCD:7641775  
Indexed Date:  
2024-03-29  
Institution Address:  
重庆邮电大学通信与信息工程学院;;信号与信息处理重庆市重点实验室, ;;信号与信息处理重庆市重点实验室, ;;, 重庆;;重庆 400065;;400065, 中国  
Institution Address:  
重庆邮电大学通信与信息工程学院;;移动通信教育部工程研究中心, ;;移动通信教育部工程研究中心, ;;, 重庆;;重庆 400065;;400065, 中国  
Institution Address:  
School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications;;Chongqing Key Laboratory of Signal and Information Processing, ;;Chongqing Key Laboratory of Signal and Information Processing, ;;, Chongqing;;Chongqing 400065;;400065  
Institution Address:  
School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications;;Engineering Research Center of Mobile Communications of the Ministry of Education, ;;Engineering Research Center of Mobile Communications of the Ministry of Education, ;;, Chongqing;;Chongqing 400065;;400065  
Institution:  
重庆邮电大学通信与信息工程学院;;信号与信息处理重庆市重点实验室  
Institution:  
重庆邮电大学通信与信息工程学院;;移动通信教育部工程研究中心  
Institution:  
School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications;;Chongqing Key Laboratory of Signal and Information Processing  
Institution:  
School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications;;Engineering Research Center of Mobile Communications of the Ministry of Education

### 446. Occluded Insulator Detection System Based on YOLOX of Multi-Scale

摘要: As a special insulation control, insulators protect transmission lines. In the process of detecting insulators by UAV, locating the insulators is a prerequisite for defect detection. However, In complex transmission line contexts, heavily occluded insulators are abundant. This kind of insulator is difficult to locate accurately due to its small size and extreme aspect ratio. To address this issue, an occlusion insulator detection system based on YOLOX is proposed in this work. Firstly, an improved SPP module is used to extract the semantic information of the backbone's P5 layer. Secondly, the multi-feature fusion module of attention mechanism (AFF-BiFPN) fully uses deep, shallow, and original feature information, extracting effective information about occluded insulators and small defects. To enhance the accuracy of insulator localization and the detection performance of the network, an adaptive anchor frame coarse extraction method is proposed. In addition, a data augmentation is proposed to simulate the occluded insulators. Experiments show that the method can detect defective insulators in the foreground while locating the occluded insulators in the background. On the test set, the insulator identification precision is 90.71%, and the recall is 88.25%.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
7  
ISSN:  
0885-8977  
eISSN:  
1937-4208  
入藏号:  
WOS:001193300900030  
Indexed Date:  
2024-04-20  
Institution Address:  
Sichuan Univ, Sch Mech Engn, Chengdu 610065, Sichuan, Peoples R China  
Institution:

### 447. Long-distance gesture detection based on deep learning for 3D spatial

摘要: With the technological development of stereoscopic display, an immersive 3D space with large size can be reconstructed more and more easily, and a 3D spatial interaction method with high-efficiency become more and more urgent. Gesture interaction, as the most natural and efficient way of human-computer interaction, can convey information very quickly and efficiently. However, the effective interaction distance of most existing gesture interaction methods is less than one meter, and can not meet the demand of the long distance 3D spatial interaction. In this paper, an efficient network named Gesture YOLO for long-distance gesture detection is proposed to achieve the small gesture object detection with improved accuracy. There are two modules in our Gesture YOLO, one is the Dual CSPDarknet53-tiny Backbone module for fusing person features and gesture features, and the other is the Progressive Multi-Scale Feature Fusion module for enhancing output features. The experimental results on our test set show that our Gesture YOLO can achieve higher gesture detection accuracy than the YOLOv4-tiny at distances ranging from 2m to 5m, and can mitigate the significant drop in gesture detection accuracy when the distance increases.  
会议名称:  
Conference on AOPC - Micro-Optics and MOEMS  
会议日期:  
JUN 20-22, 2021  
会议地点:  
Beijing, PEOPLES R CHINA  
会议赞助商:  
Univ Elect Sci & Technol China; Sci & Technol Low Light Lvel Night Vis Lab; Sci & Technol Electro Opti Informat Secur Control; Tsinghua Univ, Dept Elect Engn, Nano Optoelectron Lab; Chinese Soc Opt Engn  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0277-786X  
eISSN:  
1996-756X  
ISBN:  
978-1-5106-5008-4; 978-1-5106-5007-7  
入藏号:  
WOS:000766306400043  
Indexed Date:  
2022-05-18  
Institution Address:  
Beijing Univ Posts & Telecommun, State Key Lab Informat Photon & Opt Commun, Beijing, Peoples R China  
Institution:

### 448. TL-Detector: Lightweight Based Real-Time Traffic Light Detection Model

摘要: With the leap-forward development of intelligent driving, traffic light detection with high accuracy and high speed is important for intelligent transportation systems to ensure safety. However, existing methods encounter difficulties in balancing the detection speed and accuracy. This paper aims to present a flexible and robust lightweight model TL (Traffic Light)-Detector for real-time detection of traffic light. The model is composed of three parts. An enhanced backbone network combined with G-module is proposed to generate abundant information and reduce computational load, and the coordinate attention mechanism is introduced to focus on location features thereby strengthening the feature extract ability. The lightweight neck promotes multi-scale feature information aggregation for feature fusion to build a lightweight feature fusion network. The lightweight detection head adopts anchor-free mechanism to eliminate the hyperparameters related to the anchor. The dataset TL2022 of traffic light is built based on real traffic sceneries. Experimental results show that TL-Detector has the best comprehensive performance. TL-Detector achieves a detection speed of 277 FPS and a precision of 73.24% with only 0.72 GFLOPs. The experiments on LaRA (La Route Automatisee) public traffic light dataset show the excellent generalization ability. It indicates that TL-Detector can effectively achieve accurate and real-time detection for traffic lights.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
1524-9050  
eISSN:  
1558-0016  
入藏号:  
WOS:000980466500001  
Indexed Date:  
2023-05-17  
Institution Address:  
Sun Yat sen Univ, Sch Intelligent Syst Engn, Shenzhen 518000, Peoples R China  
Institution Address:  
Guangdong Prov Marine Engn Construct & Water Emerg, Guangzhou 510260, Peoples R China  
Institution Address:  
Guangdong Marshell Elect Technol Co, Zhaoqing 526238, Peoples R China  
Institution Address:  
Univ Sydney, Sch Aerosp Mech & Mechatron Engn, Sydney, NSW 2006, Australia  
Institution:  
Institution:  
Guangdong Prov Marine Engn Construct & Water Emerg  
Institution:  
Guangdong Marshell Elect Technol Co  
Institution:

### 449. Iterative Back Projection Network Based on Deformable 3D Convolution

摘要: Video super-resolution technology enhances the display quality of videos by obtaining high-resolution videos from low-resolution videos. Unlike single-image super-resolution, utilizing information between adjacent video frames is crucial in video super-resolution. To improve the performance of video super-resolution reconstruction, a model combining deformable 3D convolution and iterative back projection is proposed to fully exploit the temporal-spatial correlation of video frames. The model takes multiple consecutive video frames as input and outputs the super-resolution reconstruction of the middle frame, including three modules: multi-scale feature extraction, feature fusion, and high-resolution reconstruction. Firstly, multi-scale 3D convolution is used for preliminary feature extraction. Then, deformable 3D convolution and iterative back projection are combined for feature fusion. Finally, multiple residual dense blocks and sub-pixel convolution are used for high-resolution reconstruction, and global residual connections are utilized to obtain the reconstructed high-resolution video. Experimental results on the Vid4 dataset demonstrate that compared to existing methods, this method can effectively improve the peak signal-to-noise ratio and structural similarity performance and achieve better visual effects with 4x super-resolution magnification.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001101954500001  
Indexed Date:  
2023-11-27  
Institution Address:  
Xihua Univ, Coll Comp & Software Engn, Chengdu 610039, Sichuan, Peoples R China  
Institution:

### 450. MFFAE-Net: semantic segmentation of point clouds using multi-scale

摘要: Point cloud data can reflect more information about the real 3D space, which has gained increasing attention in computer vision field. But the unstructured and unordered nature of point clouds poses many challenges in their study. How to learn the global features of the point cloud in the original point cloud is a problem that has been accompanied by the research. In the research based on the structure of the encoder and decoder, many researchers focus on designing the encoder to better extract features, and do not further explore more globally representative features according to the features of the encoder and decoder. To solve this problem, we propose the MFFAE-Net method, which aims to obtain more globally representative point cloud features by using the feature learning of encoder decoder stage.Our method first enhances the feature information of the input point cloud by merging the information of its neighboring points, which is helpful for the following point cloud feature extraction work. Secondly, the channel attention module is used to further process the extracted features, so as to highlight the role of important channels in the features. Finally, we fuse features of different scales from encoding features and decoding features as well as features of the same scale, so as to obtain more global point cloud features, which will help improve the segmentation results of point clouds. Experimental results show that the method performs well on some objects in S3DIS dataset and Toronto3d dataset.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0932-8092  
eISSN:  
1432-1769  
入藏号:  
WOS:001289059900002  
Indexed Date:  
2024-08-16  
Institution Address:  
China Univ Geosci, Sch Automat, Wuhan 430074, Hubei, Peoples R China  
Institution Address:  
China Univ Geosci, Hubei Key Lab Adv Control & Intelligent Automat, Wuhan 430074, Hubei, Peoples R China  
Institution Address:  
Minist Educ, Engn Res Ctr Intelligent Technol Geoexplorat, Wuhan 430074, Hubei, Peoples R China  
Institution:  
Institution:  
Institution:  
Minist Educ

### 451. Facial Feature Extraction Algorithm Based on Improved YOLOv7-Tiny

摘要: Facial feature extraction is a critical step in driver fatigue detection, essential for improving driving safety. This paper proposes a novel driver facial feature extraction algorithm based on YOLOv7-Tiny to address the challenges of low precision and practical deployment in existing fatigue detection systems. The algorithm employs depthwise-separable convolution combined with spatial depth convolution to effectively extract key facial features. An attention module is integrated between the backbone and neck networks to enhance contextual understanding and filter out irrelevant information, enabling precise extraction of detailed and distinguishable features, even under varying lighting conditions and driver poses. Additionally, a feature fusion module is introduced to merge features from different receptive fields, improving multi-scale feature extraction and reducing the miss rate for small objects. The proposed algorithm achieves a detection accuracy of 59.8% mAP on the Drowsy-Driving-Det dataset, marking an 8.5% improvement over the original method, alongside a 60.9% reduction in model parameters. This improved algorithm not only meets real-time deployment requirements but also maintains high detection accuracy, making it well-suited for facial fatigue feature extraction in complex driving environments and edge deployment scenarios.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001422025700039  
Indexed Date:  
2025-02-21  
Institution Address:  
Xian Technol Univ, Sch Comp Sci & Engn, Xian 710021, Peoples R China  
Institution:

### 452. BGFNet: boundary information-aided graph structure fusion network for

摘要: Semantic segmentation of high-resolution remote sensing (RS) images faces challenges in multi-scale transformation. Although feature fusion is widely used in this task, the existing methods do not fully consider the spatial structure relationship between feature layers in the encoder stage. Firstly, this paper designs BGFNet network, combines the feature extraction in the encoder stage and the topological relationship modeling of graph structure, and proposes the graph structure-guided multi-scale feature fusion module to solve this problem. Secondly, in order to solve the problem of blurred object boundaries in RS image segmentation, we propose a multi-level deformable boundary guidance module, which emphasizes object boundaries by establishing long-range context. Finally, a shared enhanced attention module with shared parameters is proposed to enhance the characteristics of each class to improve the recognition ability of the model. The effectiveness of BGFNet is verified on Potsdam and Vaihingen public RS datasets, and its segmentation performance is obviously better than the existing mainstream methods. The source code will be freely available at https://github.com/zyf-cell/BGF\_Net.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001242202400001  
Indexed Date:  
2024-06-14  
Institution Address:  
Tianjin Chengjian Univ, Sch Comp & Informat Engn, Tianjin, Peoples R China  
Institution:

### 453. YOLOv8-CB: Dense Pedestrian Detection Algorithm Based on In-Vehicle

摘要: Recently, the field of vehicle-mounted visual intelligence technology has witnessed a surge of interest in pedestrian detection. Existing algorithms for dense pedestrian detection at intersections face challenges such as high computational weight, complex models that are difficult to deploy, and suboptimal detection accuracy for small targets and highly occluded pedestrians. To address these issues, this paper proposes an improved lightweight multi-scale pedestrian detection algorithm, YOLOv8-CB. The algorithm introduces a lightweight cascade fusion network, CFNet (cascade fusion network), and a CBAM attention module to improve the characterization of multi-scale feature semantics and location information, and it superimposes a bidirectional weighted feature fusion path BIFPN structure to fuse more effective features and improve pedestrian detection performance. It is experimentally verified that compared with the YOLOv8n algorithm, the accuracy of the improved model is increased by 2.4%, the number of model parameters is reduced by 6.45%, and the computational load is reduced by 6.74%. The inference time for a single image is 10.8 ms. The cascade fusion algorithm YOLOv8-CB has higher detection accuracy and is a lighter model for multi-scale pedestrian detection in complex scenes such as streets or intersections. This proposed algorithm presents a valuable approach for device-side pedestrian detection with limited computational resources.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
15  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
15  
ISSN:  
2079-9292  
入藏号:  
WOS:001139248100001  
Indexed Date:  
2024-01-19  
Institution Address:  
Shanghai Ocean Univ, Sch Engn, Shanghai 201306, Peoples R China  
Institution:

### 454. MFI-YOLO: Multi-Fault Insulator Detection Based on an Improved YOLOv8

摘要: Insulators are essential components in power transmission lines. Due to the harsh variations in bad environments, insulators may experience faults. Detecting these insulator faults promptly and effectively is an urgent issue. To rapidly and accurately locate insulators and their faulty regions in aerial images of insulators with complex backgrounds and varying fault sizes, this paper proposes an improved YOLOv8 algorithm for the detection of multiple insulator fault types (MFI-YOLO). This algorithm achieved target feature extraction in complex background images by replacing the C2F network constructed by fusing the GhostNet and multi-scale asymmetric convolution (MSA-GhostBlock). Furthermore, in the feature fusion stage, a multi-scale feature fusion structure called ResPANet, based on residual skip connections, was constructed to replace the PANet. This enhancement aims to improve the network detection accuracy in multi-target scenarios. Finally, to evaluate the algorithm's performance, this study constructed a target detection dataset containing four types of insulators: normal, self-explosive, damaged, and flashover. Experimental results indicate that, compared to the original model, the improved model has increased mean accuracy from 89.2% to 93.9%. The designed model exhibits high detection accuracy in the insulator and its three fault categories, especially for some hard-to-detect fitting.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
28  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
28  
ISSN:  
0885-8977  
eISSN:  
1937-4208  
入藏号:  
WOS:001173317300047  
Indexed Date:  
2024-04-26  
Institution Address:  
Wuhan Univ, Dept Elect Engn & Automat, Hubei Key Lab Power Equipment & Syst Secur Integra, Wuhan 430072, Peoples R China  
Institution:

### 455. A Two-Stage Density-Aware Single Image Deraining Method

摘要: Although advanced single image deraining methods have been proposed, one main challenge remains: the available methods usually perform well on specific rain patterns but can hardly deal with scenarios with dramatically different rain densities, especially when the impacts of rain streaks and the veiling effect caused by rain accumulation are heavily coupled. To tackle this challenge, we propose a two-stage density-aware single image deraining method with gated multi-scale feature fusion. In the first stage, a realistic physics model closer to real rain scenes is leveraged for initial deraining, and a network branch is also trained for rain density estimation to guide the subsequent refinement. The second stage of model-independent refinement is realized using conditional Generative Adversarial Network (cGAN), aiming to eliminate artifacts and improve the restoration quality. In particular, dilated convolutions are applied to extract rain features at multiple scales and gated feature fusion is exploited to better aggregate multi-level contextual information in both stages. Extensive experiments have been conducted on representative synthetic rain datasets and real rain scenes. Quantitative and qualitative results demonstrate the superiority of our method in terms of effectiveness and generalization ability, which outperforms the state-of-the-art.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
14  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
14  
ISSN:  
1057-7149  
eISSN:  
1941-0042  
入藏号:  
WOS:000681133600002  
PubMed ID:  
34319874  
Indexed Date:  
2021-08-11  
Institution Address:  
Wuhan Univ, Sch Elect Informat, Wuhan 430072, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Remote Sensing & Informat Engn, Wuhan 430072, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Remote Sensing & Informat Engn, Wuhan 430079, Peoples R China  
Institution Address:  
Natl Univ Singapore, Singapore Inst Neurotechnol, Singapore 117456, Singapore  
Institution Address:  
Peng Cheng Lab, Shenzhen 518055, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 456. Csb-yolo: a rapid and efficient real-time algorithm for classroom

摘要: In recent years, the integration of artificial intelligence in education has become key to enhancing the quality of teaching. This study addresses the real-time detection of student behavior in classroom environments by proposing the Classroom Student Behavior YOLO (CSB-YOLO) model. We enhance the model's multi-scale feature fusion capability using the Bidirectional Feature Pyramid Network (BiFPN). Additionally, we have designed a novel Efficient Re-parameterized Detection Head (ERD Head) to accelerate the model's inference speed and introduced Self-Calibrated Convolutions (SCConv) to compensate for any potential accuracy loss resulting from lightweight design. To further optimize performance, model pruning and knowledge distillation are utilized to reduce the model size and computational demands while maintaining accuracy. This makes CSB-YOLO suitable for deployment on low-performance classroom devices while maintaining robust detection capabilities. Tested on the classroom student behavior dataset SCB-DATASET3, the distilled and pruned CSB-YOLO, with only 0.72M parameters and 4.3 Giga Floating-point Operations Per Second (GFLOPs), maintains high accuracy and exhibits excellent real-time performance, making it particularly suitable for educational environments.  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001277780800001  
Indexed Date:  
2024-08-01  
Institution Address:  
Yunnan Normal Univ, Sch Informat Sci & Technol, Kunming 650500, Yunnan, Peoples R China  
Institution Address:  
Educ Instruments & Facil Serv Ctr, Educ Dept, Kunming 650223, Yunnan, Peoples R China  
Institution Address:  
Yunnan Normal Univ, Minist Educ, Key Lab Educ Informatizat Nationalities Minist Edu, Kunming 650500, Yunnan, Peoples R China  
Institution:  
Institution:  
Educ Instruments & Facil Serv Ctr  
Institution:

### 457. CRNet: Context feature and refined network for multi-person pose

摘要: Multi-person pose estimation is a challenging problem. Bottom-up methods have been greatly studied because the prediction speed of top-down methods is related to the number of people in the input image, making these methods difficult to apply in real-time environments. To solve the problems of scale sensitivity and quantization error in bottom-up methods, it is necessary to have a model that can predict multi-scale keypoints and refine quantization error. To achieve this, we propose context feature and refined network for multi-person pose estimation (CRNet), which can effectively solve the problems of scale sensitivity and quantization error in bottom-up methods. We use a multi-scale feature pyramid and context feature to achieve scale invariance of the network. We extract global and local features and then fuse them by attentional feature fusion (AFF) to obtain context feature that adapt to multi-scale keypoints. In addition, we propose an efficient refined network to solve the problem of quantization error and use multi-resolution supervised learning to further improve the prediction accuracy of CRNet. Comprehensive experiments are conducted on two benchmarks: COCO and MPII datasets. The average precision of CRNet reached 72.1 and 80.2%, respectively, surpassing most state-of-the-art methods.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0334-1860  
eISSN:  
2191-026X  
入藏号:  
WOS:000817903700001  
Indexed Date:  
2022-07-06  
Institution Address:  
Harbin Univ Sci & Technol, Higher Educ Key Lab Measuring & Control Technol &, Harbin 150080, Peoples R China  
Institution:

### 458. Classroom Behavior Detection Based on Improved YOLOv5 Algorithm

摘要: The detection of students' behaviors in classroom can provide a guideline for assessing the effectiveness of classroom teaching. This study proposes a classroom behavior detection algorithm using an improved object detection model (i.e., YOLOv5). First, the feature pyramid structure (FPN+PAN) in the neck network of the original YOLOv5 model is combined with a weighted bidirectional feature pyramid network (BiFPN). They are subsequently processed with feature fusion of different scales of the object to mine the fine-grained features of different behaviors. Second, a spatial and channel convolutional attention mechanism (CBAM) is added between the neck network and the prediction network to make the model focus on the object information to improve the detection accuracy. Finally, the original non-maximum suppression is improved using the distance-based intersection ratio (DIoU) to improve the discrimination of occluded objects. A series of experiments were conducted on our new established dataset which includes four types of behaviors: listening, looking down, lying down, and standing. The results demonstrated that the algorithm proposed in this study can accurately detect various student behaviors, and the accuracy was higher than that of the YOLOv5 model. By comparing the effects of student behavior detection in different scenarios, the improved algorithm had an average accuracy of 89.8% and a recall of 90.4%, both of which were better than the compared detection algorithms.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
22  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
22  
eISSN:  
2076-3417  
入藏号:  
WOS:000823512400001  
Indexed Date:  
2022-07-20  
Institution Address:  
Chongqing Normal Univ, Coll Comp & Informat Sci, Chongqing 401331, Peoples R China  
Institution Address:  
Southwest Univ, Fac Educ, Chongqing 400715, Peoples R China  
Institution:  
Institution:

### 459. SFPANet: Separation and fusion pyramid attention network for crowd

摘要: Crowd counting methods have become increasingly mature. However, the problem of dramatic scale variation still exists. For this reason, we propose an efficient separated and fused pyramid attention network, which can extract multiscale features on channels and space and greatly alleviate the problem of dramatic scale variation. First, in order to extract the rich features on the channel, we design a separated and fused channel attention module, which is composed of two 3x3 convolution layers, a separated attention module, and a SE module. Second, we design a spatial contextual feature fusion module to fully extract multiscale features in spatial dimensions. Finally, we conduct comparison experiments with state-of-the-art methods on several challenging datasets, including the ShanghaiTech, UCF\_CC\_50, and WorldExpo'10 datasets. The experimental results show our method outperforms most of the state-of-the-art methods. We conduct ablation experiments on the ShanghaiTech Part A and Part B datasets to verify the importance of each submodule.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001081680200011  
Indexed Date:  
2023-11-06  
Institution Address:  
East China Jiaotong Univ, Sch Informat Engn, Shuanggang East St, Nanchang 330013, Jiangxi, Peoples R China  
Institution Address:  
Jiangxi Prov Commun Investment Grp Co Ltd, Rd Network Operat Management Co, Shuanggang East St, Nanchang 330013, Jiangxi, Peoples R China  
Institution Address:  
Jiangxi Univ Finance & Econ, Sch Software & Internet Things Engn, Shuanggang East St, Nanchang 330013, Jiangxi, Peoples R China  
Institution:  
Institution:  
Jiangxi Prov Commun Investment Grp Co Ltd  
Institution:

### 460. AFRE-Net: Adaptive Feature Representation Enhancement for Arbitrary

摘要: Arbitrary-oriented object detection (AOOD) is a crucial task in aerial image analysis but is also faced with significant challenges. In current AOOD detectors, commonly used multi-scale feature fusion modules fall short in spatial and semantic information complement between scales. Additionally, fixed feature extraction structures are usually used following a fusion model, resulting in the inability of detectors to self-adjust. At the same time, feature fusion and extraction modules are designed in isolation and the internal synergy between them is ignored. The above problems result in feature representation deficiency, thus affecting the overall detection precision. To solve these problems, we first create a fine-grained feature pyramid network (FG-FPN) that not only provides richer spatial and semantic features, but also completes neighbor scale features in a self-learning mode. Subsequently, we propose a novel feature enhancement module (FEM) to fit FG-FPN. FEM authorizes the detection unit to automatically adjust the sensing area and adaptively suppress background interference, thereby generating stronger feature representations. Our proposed solution was tested through extensive experiments on challenging datasets, including DOTA (77.44% mAP), HRSC2016 (97.82% mAP), UCAS-AOD (91.34% mAP), as well as ICDAR2015 (86.27% F-score) and its effectiveness and high applicability are verified on all the above datasets.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:001089379200001  
Indexed Date:  
2023-11-05  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Key Lab Computat Opt Imaging Technol, Beijing 100094, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Key Lab Digital Earth Sci, Beijing 100094, Peoples R China  
Institution Address:  
Chinese Acad Sci, Int Res Ctr Big Data Sustainable Dev Goals, Beijing 100094, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Coll Resources & Environm, Beijing 100049, Peoples R China  
Institution Address:  
Univ Tokyo, Dept Complex Sci & Engn, Tokyo 2778561, Japan  
Institution Address:  
Harbin Engn Univ, Coll Intelligent Syst Sci & Engn, Harbin 150001, Peoples R China  
Institution Address:  
Harbin Engn Univ, Qingdao Innovat & Dev Ctr, Qingdao 266000, Peoples R China  
Institution Address:  
Liaocheng Univ, Coll Geog & Environm, Liaocheng 252059, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 461. Pyramid-attention based multi-scale feature fusion network for

摘要: Remote sensing images with high spatial resolution and high spectral resolution have important applications in human society. In general, due to the limitations faced by the optical sensors', we are limited to obtain only low spatial resolution multispectral images (MS) and high spatial resolution panchromatic images (PAN). To address this limitation, this study proposes a pyramid-attention based multi-scale feature fusion network (PAMF-Net) that combines the pyramid attention mechanism and feature aggregation. Initially, the MS and PAN images are input to the network, and the PAN images pass through the input pyramid branch to generate a multi-level receiving domain. Then, the result is combined with the features of the MS image as the input of the encoder, and these composite features are input to the pyramid attention mechanism module to capture multi-scale corresponding features. Next, the result of the input pyramid branch is input to the feature aggregation module to seamlessly merge with the features of the pyramid attention mechanism. Finally, in the encoding stage, multiple levels of features are multiplexed as encoding secondary lines by skipping connections to obtain high-quality HRMS images. After quantitative and qualitative experiments, the results show that our method is superior to other advanced methods.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
11  
被引频次合计:  
11  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000683627400003  
Indexed Date:  
2021-08-21  
Institution Address:  
Shandong Technol & Business Univ, Sch Informat & Elect Engn, Yantai 264005, Peoples R China  
Institution Address:  
Shandong Technol & Business Univ, Sch Comp Sci & Technol, Yantai 264005, Peoples R China  
Institution:  
Institution:

### 462. Enhanced YOLOv7 for Improved Underwater Target Detection

摘要: Aiming at the problems of the underwater existence of some targets with relatively small size, low contrast, and a lot of surrounding interference information, which lead to a high leakage rate and low recognition accuracy, a new improved YOLOv7 underwater target detection algorithm is proposed. First, the original YOLOv7 anchor frame information is updated by the K-Means algorithm to generate anchor frame sizes and ratios suitable for the underwater target dataset; second, we use the PConv (Partial Convolution) module instead of part of the standard convolution in the multi-scale feature fusion module to reduce the amount of computation and number of parameters, thus improving the detection speed; then, the existing CIou loss function is improved with the ShapeIou\_NWD loss function, and the new loss function allows the model to learn more feature information during the training process; finally, we introduce the SimAM attention mechanism after the multi-scale feature fusion module to increase attention to the small feature information, which improves the detection accuracy. This method achieves an average accuracy of 85.7% on the marine organisms dataset, and the detection speed reaches 122.9 frames/s, which reduces the number of parameters by 21% and the amount of computation by 26% compared with the original YOLOv7 algorithm. The experimental results show that the improved algorithm has a great improvement in detection speed and accuracy.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
eISSN:  
2077-1312  
入藏号:  
WOS:001277270700001  
Indexed Date:  
2024-08-02  
Institution Address:  
Jiangsu Univ Sci & Technol, Sch Mech Engn, Zhenjiang 212003, Peoples R China  
Institution Address:  
Jiangsu Univ Sci & Technol, Marine Equipment & Technol Inst, Zhenjiang 212003, Peoples R China  
Institution:  
Institution:

### 463. SA-GCN: structure-aware graph convolutional networks for crowd pose

摘要: In this paper, we aim to capture the structure dependency of human joints and improve the localization accuracy of invisible joints. We propose a novel framework: Structure-aware Graph Convolutional Network (SA-GCN) for crowd pose estimation, which can be divided into two components: Sample Pose Net and Refined Pose Net. Firstly, Sample Pose Net includes a multi-scale feature fusion module, which uses multi-scale features to capture small-scale characters and extract the global "rough" pose as much as possible. Secondly, channel and spatial attention are injected into the multi-scale feature fusion module to strengthen the characteristics of small-scale characters. Finally, graph convolution obtained by the disentangled several parallel sub-graph convolution modules in Refined Pose Net. Global and structural advantages of graph convolution are more conducive to predicting difficult points in sample Pose. In addition, SA-GCN obtains lower parameters compared with the popular pose estimation networks. By which, we apply a novel framework SA-GCN to get feature maps for proposal and refinement, respectively. Comprehensive experiments demonstrate that the proposed method achieves superior pose estimation results on two benchmark datasets, CrowdPose and MSCOCO. Moreover, SA-GCN significantly outperforms state-of-the-art performance on CrowdPose and almost always generates plausible human pose predictions.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0920-8542  
eISSN:  
1573-0484  
入藏号:  
WOS:000923438700002  
Indexed Date:  
2023-02-20  
Institution Address:  
Huaqiao Univ, Coll Comp Sci & Technol, Xiamen 361021, Peoples R China  
Institution Address:  
Huaqiao Univ, Xiamen Key Lab Comp Vis & Pattern Recognit, Xiamen 361021, Peoples R China  
Institution:  
Institution:

### 464. EFLightCaps: an efficient feature-focused lightweight capsule network

摘要: The prevalent issues of small samples and heavy noise in industrial settings have severely limited the application effectiveness of intelligent diagnostic methods. To address these challenges, an efficient feature-focused lightweight capsule network is proposed. Firstly, a plug-and-play lightweight Ghost multiscale convolution block is designed to integrate multiscale convolution with gating mechanisms through a multi-branch architecture, enabling multi-scale feature fusion and enhancing the model's noise robustness. Secondly, an efficient feature-focused routing mechanism is proposed to optimize information transfer between capsules through multi-step feature reconstruction and feature focusing strategies, which substantially enhances generalization capabilities in small-sample scenarios. Finally, a dual-norm frequency-domain regularization loss function is designed to leverage the complementary advantages of first-order and second-order norms, enhancing both feature extraction capability and reconstruction quality. Extensive experiments on two distinct datasets demonstrate that EFLightCaps achieves superior diagnostic performance and computational efficiency in scenarios with small samples and heavy noise.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001427554400001  
Indexed Date:  
2025-02-26  
Institution Address:  
Army Mil Transportat Univ, Dept Mil Vehicle Engn, Tianjin 300161, Peoples R China  
Institution Address:  
Tianjin Univ, Sch Mech Engn, Tianjin 300350, Peoples R China  
Institution:  
Army Mil Transportat Univ  
Institution:

### 465. Attention-based fusion factor in FPN for object detection

摘要: At present, most advanced detectors usually use the feature pyramid to detect objects of different scales. Among them, FPN is one of the representative works of multi-scale feature summation to construct the feature pyramid. However, the existing FPN-based feature extraction networks pay more attention to capturing effective semantic information and ignore the influence of the dataset scale distribution on the FPN feature fusion process. To solve this problem, we propose a novel attention structure, which can be applied to any FPN-based network model. Different from the general attention that gets its own attention from itself, our proposed method makes better use of the influence of the lower layer feature of the adjacent layer on feature fusion, which guides the filtering of the upper layer feature. By considering the difference in the feature information of the same sample in different feature maps, it is better to filter out the invalid sample features of the upper layer relative to the lower layer. Our method can better learn the degree of deep features participating in shallow learning so that each layer of FPN is more focused on its own layer learning while effectively transferring features. Our experimental results show that our method can significantly improve the multi-scale object detection performance of the model.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
17  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
4  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
21  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000769862200008  
Indexed Date:  
2022-03-31  
Institution Address:  
North China Elect Power Univ, Sch Control & Comp Engn, Beijing, Peoples R China  
Institution:

### 466. Gradient-Guided and Multi-Scale Feature Network for Image

摘要: Recently, deep-learning-based image super-resolution methods have made remarkable progress. However, most of these methods do not fully exploit the structural feature of the input image, as well as the intermediate features from the intermediate layers, which hinders the ability of detail recovery. To deal with this issue, we propose a gradient-guided and multi-scale feature network for image super-resolution (GFSR). Specifically, a dual-branch structure network is proposed, including the trunk branch and the gradient one, where the latter is used to extract the gradient feature map as structural prior to guide the image reconstruction process. Then, to absorb features from different layers, two effective multi-scale feature extraction modules, namely residual of residual inception block (RRIB) and residual of residual receptive field block (RRRFB), are proposed and embedded in different network layers. In our RRIB and RRRFB structures, an adaptive weighted residual feature fusion block (RFFB) is investigated to fuse the intermediate features to generate more beneficial representations, and an adaptive channel attention block (ACAB) is introduced to effectively explore the dependencies between channel features to further boost the feature representation capacity. Experimental results on several benchmark datasets demonstrate that our method achieves superior performance against state-of-the-art methods in terms of both subjective visual quality and objective quantitative metrics.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
eISSN:  
2076-3417  
入藏号:  
WOS:000776008500001  
Indexed Date:  
2022-04-07  
Institution Address:  
Huaqiao Univ, Coll Engn, Quanzhou 362021, Peoples R China  
Institution:

### 467. Garbage image classification algorithm based on improved MobileNet v2

摘要: A garbage image classification method based on improved MobileNet v2 was proposed aiming at the problems of poor real-time performance and low classification accuracy of existing garbage image classification models. A lightweight feature extraction network based on MobileNet v2 was constructed. The parameter numbers of the model were reduced by adjusting its width factor, channel and spatial attention modules were embedded in the model to enhance the network's ability to refine features, a multi-scale feature fusion structure was designed to enhance the adaptability of the network to scale, and transfer learning was used to optimize the model parameters to further improve the model accuracy. Experimental results show that the average accuracy of the algorithm on the self built dataset was 94.6%, which was 2.0%, 3.4%, 3.2%, 2.3% and 1.2% higher than that of MobileNet v2, VGG16, GoogleNet, ResNet50 and ResNet101 models, respectively. The proposed algorithm achieved good performance in two public image classification datasets, CIFAR-100 and tiny-ImageNet. The parameter numbers of the model was only 0.83 M, which was about 2/5 of the basic model. The single inference on edge device JETSON TX2 took 68 ms, which proved the improvement of inference speed and prediction accuracy.  
摘要:  
针对现有的垃圾图像分类模型实时性能差和分类精度低的问题,提出基于改进MobileNet v2的垃圾图像分类方法,构建以MobileNet v2为核心的轻量级特征提取网络.通过调整宽度因子降低模型的参数量;在模型中嵌入通道和空间注意力模块,增强网络对特征的细化能力;设计多尺度特征融合结构,增强网络对尺度的适应性;利用迁移学习的方式优化模型参数,进一步提高模型精度.实验结果表明,算法在自建数据集上的平均准确率为94.6%,分别高于MobileNet v2、VGG16、GoogleNet、ResNet50、ResNet101模型2.0%、3.4%、3.2%、2.3%、1.2%;所提算法在2种公共图像分类数据集CIFAR-100和tiny-ImageNet中均取得不错表现;模型参数量仅为0.83 M,体积约为基础模型的2/5,在边缘设备JETSON TX2上的单次推理耗时68 ms,实现了推理速度和预测准确率的提升.  
Web of Science 核心合集中的 "被引频次":  
8  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
1008-973X  
入藏号:  
CSCD:7044158  
Indexed Date:  
2021-11-19  
Institution Address:  
江西理工大学电气工程与自动化学院;;江西省磁悬浮技术重点实验室, ;;江西省磁悬浮技术重点实验室, 赣州;;赣州, 江西;;江西 341000;;341000, 中国  
Institution Address:  
School of Electrical Engineering and Automation, Jiangxi University of Science and Technology;;Jiangxi Provincial Key Laboratory of Maglev Technology, ;;Jiangxi Provincial Key Laboratory of Maglev Technology, Ganzhou;;Ganzhou, ;; 341000;;341000  
Institution:  
江西理工大学电气工程与自动化学院;;江西省磁悬浮技术重点实验室  
Institution:  
School of Electrical Engineering and Automation, Jiangxi University of Science and Technology;;Jiangxi Provincial Key Laboratory of Maglev Technology

### 468. Synergistic Attention Network for RGB-D Saliency Detection

摘要: RGB information and depth information are both crucial feature cues in RGB-D saliency detection. However, the current RGB-D saliency detection models are unable to handle the interaction of multi-scale features and the fusion of multi-modal features efficiently, so they are limited in understanding natural scenes. To alleviate such shortcomings, we propose a synergistic attention network for RGB-D saliency detection (SANet), and introduce a general multi-modal learning strategy, termed Guidance and Teaching Strategy. The implicit guidance strategy works in the multi-scale feature extraction stage of the encoder, while the explicit teaching strategy works in the feature fusion stage of the decoder, so as to realize the phased learning process of encoding and decoding stages. Extensive experiments are conducted on four saliency detection benchmarks, and the results verify the proposed method performs better than the 18 cutting-edge RGB-D saliency detection models in terms of four metrics.  
摘要:  
在RGB-D显著性检测视觉任务中,RGB彩色模态和深度模态的信息均被视为十分重要的特征线索。但现有的RGB-D显著性检测模型无法高效执行多尺度特征的交互和多模态特征的融合,因此在真实的开放场景下表现欠佳。针对上述问题,提出了一种基于协同注意力(synergistic attention)机制的RGB-D显著性检测算法模型(SANet),并引入多模态学习中通用的引导与教导策略(guidance and teaching strategy)。在编码器进行多尺度特征提取的阶段中进行隐式引导(implicit guidance),在解码器进行特征融合时进行显式的教导(explicit teaching),实现了编码、解码的分阶段学习。在4个显著性检测评测数据集上进行的综合实验表明,该算法在4个评测指标上均优于已有的18个前沿RGB-D显著性检测模型。  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1671-8836  
入藏号:  
CSCD:7411419  
Indexed Date:  
2023-06-10  
Institution Address:  
武汉大学计算机学院, 武汉, 湖北 430072, 中国  
Institution Address:  
武汉船舶通信研究所, 武汉, 湖北 430205, 中国  
Institution Address:  
School of Computer Science, Wuhan University, Wuhan, Hubei 430072, China  
Institution Address:  
Wuhan Maritime Communication Research Institute, Wuhan, Hubei 430205, China  
Institution:  
武汉大学计算机学院  
Institution:  
武汉船舶通信研究所  
Institution:  
School of Computer Science, Wuhan University  
Institution:  
Wuhan Maritime Communication Research Institute

### 469. Cross-level interaction fusion network-based RGB-T semantic segmentation

摘要: RGB-T segmentation represents an innovative approach driven by advancements in multispectral detection and is poised to replace traditional RGB segmentation methods. An effective cross-modality feature fusion module is essential for this technology. The precise segmentation of distant objects is another significant challenge. Focused on these two areas, we propose an end-to-end distant object feature fusion network (DOFFNet) for RGB-T segmentation. Initially, we introduce a cross-level interaction fusion strategy (CLIF) and an inter-correlation fusion method (IFFM) in the encoder to enhance multi-scale feature expression and improve fusion accuracy. Subsequently, we propose a residual dense pixel convolution (R-DPC) in the decoder with a trainable upsampling unit that dynamically reconstructs information lost during encoding, particularly for distant objects whose features may vanish after pooling. Experimental results show that our DOFFNet achieves a top mean pixel accuracy of 75.8% and dramatically improves accuracy for four classes, including objects occupying as little as 0.2%-2% of total pixels. This improvement ensures more reliable and effective performance in practical applications, particularly in scenarios where small object detection is critical. Moreover, it demonstrates potential applicability in other fields like medical imaging and remote sensing.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:001388431800001  
Indexed Date:  
2025-01-07  
Institution Address:  
Xidian Univ, Sch Optoelect Engn, Xian 710071, Peoples R China  
Institution Address:  
Beijing Inst Control & Elect Technol, Beijing 100038, Peoples R China  
Institution Address:  
Natl Key Lab Sci & Technol Test Phys & Numer Math, Beijing 100076, Peoples R China  
Institution:  
Institution:  
Beijing Inst Control & Elect Technol  
Institution:  
Natl Key Lab Sci & Technol Test Phys & Numer Math

### 470. Coupled Global-Local object detection for large VHR aerial images

摘要: Object detection in large aerial images generally requires splitting each image into local images in the preprocessing step, and even the state-of-the-art models currently use this preprocessing method. However, image splitting often leads to deficiencies in contextual information and incomplete detec-tion of oversized objects. At present, many object detection methods are designed to deal with large images. However, they require complex additional structures or training steps, and their applicability is limited. To address these problems, we propose the Coupled Global-Local (CGL) network, which can be easily embedded in frequently used detection models, to efficiently capture more information. Specifically, we employ a multiscale feature fusion module to share information between the global and local branches. Furthermore, a new convolution method is proposed to adaptively adjust the receptive field for better feature extraction. In addition, we find that detection results from global branches in the existing global-local architecture hinder the performance improvement on details when the detection results from different-resolution branches are fused. Therefore, on the global branch, a proposal filter and a nonlocal suppression (NLS) algorithm are developed to prevent small positive proposals and remove unqualified detection boxes easily and efficiently, respectively. We conduct extensive experiments on the DOTA-1.0, DOTA-1.5, and DOTA-2.0 data sets. The results demonstrate that CGL can significantly improve the detection performance of various baseline models for large very-high-resolution (VHR) aerial images without bells and whistles.(c) 2022 Elsevier B.V. All rights reserved.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
14  
ISSN:  
0950-7051  
eISSN:  
1872-7409  
入藏号:  
WOS:000992243600001  
Indexed Date:  
2023-06-07  
Institution Address:  
East China Normal Univ, Key Lab Geog Informat Sci, Minist Educ, Shanghai 200241, Peoples R China  
Institution Address:  
East China Normal Univ, Sch Geog Sci, Shanghai 200241, Peoples R China  
Institution Address:  
Minist Nat Resources, Key Lab Spatial Temporal Big Data Anal & Applicat, Shanghai 200241, Peoples R China  
Institution Address:  
East China Normal Univ, Shanghai Key Lab Multidimens Informat Proc, Shanghai 200241, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Comp Sci & Technol, Beijing 101408, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Shenzhen Campus, Shenzhen, Guangdong, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 471. A multimodal fusion framework for urban scene understanding and

摘要: Urban scene understanding and functional identification are essential for accurately characterizing the spatial structure and optimizing the city layouts during rapid urbanization. Multimodal data is important for recognizing the distribution patterns of urban functions and revealing internal details. Previous studies have focused primarily on remote sensing imagery and points of interest (POIs) data, overlooking the role of building characteristics in determining functions of urban scenes. These studies are also limited in terms of mining and fusing multimodal features. To address these challenges, this study proposes a multimodal fusion framework that integrates remote sensing imagery, POIs, and building footprints for urban scene understanding and functional mapping. The framework employs a dual-branch model that extracts visual semantic features from the remote sensing imagery and socioeconomic features from auxiliary data, such as POIs and building footprints. A branch attention module is designed to assign weights to dual-branch features. Additionally, a multiscale feature fusion module is introduced to extract and combine multiscale features through modal interaction. Experiments in Beijing and Chengdu validate the effectiveness of the proposed framework with overall accuracy of 90.04% and 92.07%, and kappa coefficient of 0.881 and 0.895, respectively. This study provides empirical evidence to support accurate urban planning and further promote urban sustainable development. The source code is at: htt ps://github.com/sssuchen/MMFF.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
11  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
1569-8432  
eISSN:  
1872-826X  
入藏号:  
WOS:001178157800001  
Indexed Date:  
2024-03-27  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution Address:  
Hainan Aerosp Informat Res Inst, Key Lab Earth Observat Hainan Prov, Sanya 572029, Peoples R China  
Institution:  
Institution:  
Institution:  
Hainan Aerosp Informat Res Inst

### 472. Faster object detector for drone-captured images

摘要: Object detection based on computer vision is becoming popular in drone-captured images. However, real-time object detection in unmanned aerial vehicle (UAV) scenarios is a huge challenge for low-end devices. To deal with the problem, we have improved YOLOv3-tiny in the following aspects. First, the label rewriting problem, which is caused by network structure and dataset of YOLOv3-tiny in drone-captured images detection, is very serious. The method of increasing the size of the predicted feature map is used to reduce the ratio of label rewriting. Second, the features of small targets will be reduced in a small feature map, but the context information with large receptive fields in it can improve the performance of small target detection. So we use dilated convolution to expand the receptive field without reducing the size of the feature map. Third, multiscale feature fusion is very helpful for small target detection. The multidilated module is adopted to merge features in earlier layer and deeper layers. Finally, a pretraining strategy combining copy-paste data augmentation method is proposed to learn more features from categories with a small number of samples. We evaluated our model on the VisDrone2019-Det test set. It achieves compelling results compared to the counterparts of YOLOv3-tiny, including similar to 86.1 % decline in model size, increasing similar to 19.2 % AP50. Although our model is slower than YOLOv3-tiny, it is 2.96 times faster than YOLOv3. The results of experiments verify that our network is more effective than YOLOv3-tiny. It is more suitable for UAV object detection applications on low-end devices.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:000848751400051  
Indexed Date:  
2022-09-09  
Institution Address:  
Southwest Petr Univ, Sch Comp Sci, Chengdu, Peoples R China  
Institution Address:  
Univ Elect Sci & Technol China, Sch Informat & Commun Engn, Chengdu, Peoples R China  
Institution Address:  
Sichuan Police Coll, Dept Traff Management, Luzhou, Peoples R China  
Institution:  
Institution:  
Institution:

### 473. Flexible asymmetric convolutional attention network for LiDAR semantic

摘要: LiDAR semantic segmentation is an essential task in understanding 3D semantic information. Currently, the most efficient approach to LiDAR data segmentation is to project the point cloud into the 2D plane and process it using 2D convolution. The results of this approach are encouraging. However, the elevation angle of LiDAR is larger than the azimuth angle, resulting in the range map being vertically elongated in the 3D space captured per unit pixel area. If a square convolution kernel is used, the extracted features will be distorted. To address these limitations, we propose the flexible asymmetric convolutional attention network (FACANet), built from flexible asymmetric convolution and lightweight decoding modules. In this encoder structure, a meta-kernel accounts for the geometric information in 3D space, which helps encode the input range image features effectively. Moreover, a flexible asymmetric convolutional attention block (FACAB) is proposed to capture elongated features in the range image. To facilitate lightweight decoding, the channel uniform interpolation block (CUIB) uses 1x1\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$1\times 1$$\end{document} convolutions to reduce channels and bilinear interpolation to upsample features at each resolution. Furthermore, the continuous multiscale feature fusion block (CMFB) is proposed to fuse features at different resolutions. Finally, a convolutional spatial propagation network (CSPN)-based segmentation head is introduced to improve the accuracy of the segmentation results. Quantitative and qualitative experiments are conducted on the public datasets SemanticKITTI and SemanticPOSS, and our approach achieves better accuracy than advanced models.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:001230097700001  
Indexed Date:  
2024-05-30  
Institution Address:  
China Univ Min & Technol Beijing, Sch Artificial Intelligence, Coll Rd St, Beijing 100083, Peoples R China  
Institution:

### 474. Semantic Segmentation for Road Scene Based on Multiscale Feature Fusion

摘要: A lightweight network model based on multiscale feature information fusion (MIFNet) is developed in this study owing to the imbalance among the parameter amount, inference speed, and accuracy in many existing semantic segmentation network models. The MIFNet is constructed on the encoding-decoding architecture. In the encoding part, the split strategy and asymmetric convolution are flexibly applied to design lightweight bottleneck structure for feature extraction. The spatial attention mechanism and Laplace edge detection operator are introduced to fuse spatial and edge information to obtain rich feature information. In the decoding part, a new decoder is designed by introducing a channel attention mechanism to recover the size and detail information of the feature map for a complete semantic segmentation task. The MIFNet achieves accuracies of 73. 1% and 67. 7% on the Cityscapes and CamVid test sets, respectively, with only approximately 0. 82 M parameters. Correspondingly, it reaches up to 73. 68 frame/s and 85. 16 frame/s inference speed, respectively using a single GTX 1080Ti GPU. The results show that the method achieves a good balance in terms of the parameter amount, inference speed, and accuracy, yielding a lightweight, fast, and accurate semantic segmentation.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1006-4125  
入藏号:  
WOS:001083490000004  
Indexed Date:  
2023-11-07  
Institution Address:  
Jinan Univ, Coll Informat Sci & Technol, Guangzhou 510632, Guangdong, Peoples R China  
Institution:

### 475. Multi-scale feature fusion with attention mechanism for crowded road

摘要: Crowded object detection under the heavy traffic environment is always a challenging task in the field of autonomous driving and robotics, because the dense gathering of vehicles or pedestrians inevitably bring heavy occlusion. It is difficult to distinguish highly overlapped objects and predict their bounding boxes accurately, especially for small objects far down the road. To address this challenge, this paper proposes an improved YOLOv5s network integrating a multi-scale feature fusion module with attention mechanism for crowded road object detection task. Specifically, to enhance the multi-scale representation of semantic features and to model the object scale variation flexibly, we introduce an attention-guided pyramid feature fusion strategy into the YOLOv5s backbone network. Then a C3CA module is designed by embedding the coordinate attention (CA) into the concentrated-comprehensive convolution (C3) module of the original YOLOv5s, which can boost the ability of extracting distinguishing features from the overlapped objects. In addition, we add implicit detection heads (IDHs) into the original YOLOv5s's detection head part, which helps the network to learn implicit knowledge and improves the detection accuracy. Finally, a simplified optimal transport assignment (SimOTA) and a bounding box regression loss with dynamic focusing mechanism are used to improve the detector's overall performance. Extensive experiments on the public dataset BDD100K and our self-built crowded road object dataset (XMRD) demonstrate the superiority of our model in crowded road scenarios. The mean average precision (mAP) of our model can achieve 71.2% and 88.2% on the BDD100K and XMRD datasets, respectively, which provides an improvement of +3% over the existing state of the art models.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001159319900001  
Indexed Date:  
2024-02-17  
Institution Address:  
Hangzhou Dianzi Univ, Sch Comp Sci & Technol, Hangzhou, Peoples R China  
Institution Address:  
Hangzhou Xinmai Silicon Technol Inc, Hangzhou, Peoples R China  
Institution:  
Institution:  
Hangzhou Xinmai Silicon Technol Inc

### 476. Multi-Stage Salient Object Detection in 360° Omnidirectional Image Using

摘要: Recently, salient object detection (SOD) for 2D images has been extensively studied. However, due to the complexity of scene and the existence of geometric distortions, research on 360 degrees SOD is still lacking with respect to the wide field-of-view. In this paper, we explore a multi-stage solution for SOD of 360 degrees omnidirectional images, which considers the effects of RGB image and the complementary object-level semantic (OLS) information in locating the objects. Specifically, to effectively concatenate two types of features, we propose a novel Multi-level Feature Fusion and Progressive Aggregation Network (MFFPANet) for accurately detecting the salient objects in 360 degrees omnidirectional images, which is mainly composed of a dynamic complementary feature fusion (DCFF) module and a progressive multi-scale feature aggregation (PMFA) module. First, the OLS and RGB images share the same backbone network for joint learning, and the DCFF module dynamically integrates the hierarchical features from the backbone network. In addition, the PMFA module includes multiple cascaded feature integration modules, which gradually integrate multi-scale features via deep supervision in a progressive manner. Experimental results show that the proposed MFFPANet achieves superior performances on two 360 degrees SOD databases.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
2471-285X  
入藏号:  
WOS:001005996100001  
Indexed Date:  
2023-06-24  
Institution Address:  
Ningbo Univ, Fac Informat Sci & Engn, Ningbo 315211, Peoples R China  
Institution Address:  
Gwangju Inst Sci & Technol, Sch Informat & Commun, Gwangju 500712, South Korea  
Institution:  
Institution:

### 477. TSBA-YOLO: An Improved Tea Diseases Detection Model Based on Attention

摘要: Tea diseases have a significant impact on the yield and quality of tea during the growth of tea trees. The shape and scale of tea diseases are variable, and the tea disease targets are usually small, with the intelligent detection processes of tea diseases also easily disturbed by the complex background of the growing region. In addition, some tea diseases are concentrated in the entire area of the leaves, needing to be inferred from global information. Common target detection models are difficult to solve these problems. Therefore, we proposed an improved tea disease detection model called TSBA-YOLO. We use the dataset of tea diseases collected at the Maoshan Tea Factory in China. The self-attention mechanism was used to enhance the ability of the model to obtain global information on tea diseases. The BiFPN feature fusion network and adaptively spatial feature fusion (ASFF) technology were used to improve the multiscale feature fusion of tea diseases and enhance the ability of the model to resist complex background interference. We integrated the Shuffle Attention mechanism to solve the problem of difficult identifications of small-target tea diseases. In addition, we used data-enhancement methods and transfer learning to expand the dataset and relocate the parameters learned from other plant disease datasets to enhance tea diseases detection. Finally, SIoU was used to further improve the accuracy of the regression. The experimental results show that the proposed model is good at solving a series of problems encountered in the intelligent recognition of tea diseases. The detection accuracy is ahead of the mainstream target detection models, and the detection speed reaches the real-time level.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
12  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
29  
被引频次合计:  
31  
eISSN:  
1999-4907  
入藏号:  
WOS:000958573200001  
Indexed Date:  
2023-04-14  
Institution Address:  
Nanjing Forestry Univ, Coll Informat Sci & Technol, Nanjing 210037, Peoples R China  
Institution Address:  
Nanjing Agr Univ, Coll Informat Management, Nanjing 210095, Peoples R China  
Institution Address:  
McMaster Univ, Dept Comp & Software, Hamilton, ON L8S 4L8, Canada  
Institution:  
Institution:  
Institution:

### 478. A Study of English Learning Vocabulary Detection Based on Image Semantic

摘要: College students learn words always under both teachers' and school administrators' control. Based on multi-modal discourse analysis theory, the analysis of English words under the synergy of different modalities, students improve the motivation and effectiveness of word learning, but there are still some problems, such as the lack of visual modal memory of pictures, incomplete word meanings, little interaction between users, and lack of resource expansion function. To this end, this paper proposes a stepped image semantic segmentation network structure based on multi-scale feature fusion and boundary optimization. The network aims at improving the accuracy of the network model, optimizing the spatial pooling pyramid module in Deeplab V3+ network, using a new activation function Funnel ReLU (FReLU) for vision tasks to replace the original non-linear activation function to obtain accuracy compensation, improving the overall image segmentation accuracy through accurate prediction of the boundaries of each class, reducing the intra-class error in the prediction results. The accuracy compensation is obtained by replacing the original linear activation function with FReLU. Experimental results on the Englishhnd dataset demonstrate that the improved network can achieve 96.35% accuracy for English characters with the same network parameters, training data and test data.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
1662-5188  
入藏号:  
WOS:000812001900001  
PubMed ID:  
35720773  
Indexed Date:  
2022-06-24  
Institution Address:  
Zhejiang Business Coll, Sch Int Studies, Hangzhou, Peoples R China  
Institution:  
Zhejiang Business Coll

### 479. Target detection for remote sensing based on the enhanced YOLOv4 with

摘要: To solve problems for false detection, inadequate regression performance of anchor frames, and the inability to detect small targets in traditional multiscale target detection methods based on YOLOv4, we propose a novel target detection framework named as Enhanced YOLOv4. Firstly, our improved BiFPN replaced the original PANet as the feature fusion module, which can achieve multi-scale feature fusion by way of shared weights. Secondly, the channel attention mechanism (CAM) was embedded before the detection head to highlight the correlation between channels so that small targets can be get more attention. At last, to improve the anchor box regression effect and accelerate the training speed of YOLOv4, we improved the net training loss function, in which the original CIoU was replaced by CDIoU. The experimental results on the DOTA dataset validate our improvement. The mAP of our method is 90.88%, and the frame rate reached 58.76 FPS, at the same time, the speed of detection is not affected significantly.(c) 2023 National Authority of Remote Sensing & Space Science. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
15  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
17  
ISSN:  
1110-9823  
eISSN:  
2090-2476  
入藏号:  
WOS:001002493600001  
Indexed Date:  
2023-06-17  
Institution Address:  
Heilongjiang Univ, Elect Engn Coll, Key Lab Remote Sensing Image Proc, Harbin 150080, Peoples R China  
Institution:

### 480. Research on Rejoining Bone Stick Fragment Images: A Method Based on

摘要: The rejoining of bone sticks holds significant importance in studying the historical and cultural aspects of the Han Dynasty. Currently, the rejoining work of bone inscriptions heavily relies on manual efforts by experts, demanding a considerable amount of time and energy. This paper introduces a multi-scale feature fusion Siamese network guided by edge contour (MFS-GC) model. Constructing a Siamese network framework, it first uses a residual network to extract features of bone sticks, which is followed by computing the L2 distance for similarity measurement. During the extraction of feature vectors using the residual network, the BN layer tends to lose contour detail information, resulting in less conspicuous feature extraction, especially along fractured edges. To address this issue, the Spatially Adaptive DEnormalization (SPADE) model is employed to guide the normalization of contour images of bone sticks. This ensures that the network can learn multi-scale boundary contour features at each layer. Finally, the extracted multi-scale fused features undergo similarity measurement for local matching of bone stick fragment images. Additionally, a Conjugable Bone Stick Dataset (CBSD) is constructed. In the experimental validation phase, the MFS-GC algorithm is compared with classical similarity calculation methods in terms of precision, recall, and miss detection rate. The experiments demonstrate that the MFS-GC algorithm achieves an average accuracy of 95.5% in the Top-15 on the CBSD. The findings of this research can contribute to solving the rejoining issues of bone sticks.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
2076-3417  
入藏号:  
WOS:001149478600001  
Indexed Date:  
2024-02-04  
Institution Address:  
Xian Univ Architecture & Technol, Sch Informat & Control Engn, Xian 710055, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Archaeol, Beijing 100101, Peoples R China  
Institution Address:  
Shaanxi Inst Preservat Cultural Heritage, Xian 710075, Peoples R China  
Institution:  
Institution:  
Institution:  
Shaanxi Inst Preservat Cultural Heritage

### 481. FFU-Net: Feature Fusion U-Net for Lesion Segmentation of Diabetic

摘要: Diabetic retinopathy is one of the main causes of blindness in human eyes, and lesion segmentation is an important basic work for the diagnosis of diabetic retinopathy. Due to the small lesion areas scattered in fundus images, it is laborious to segment the lesion of diabetic retinopathy effectively with the existing U-Net model. In this paper, we proposed a new lesion segmentation model named FFU-Net (Feature Fusion U-Net) that enhances U-Net from the following points. Firstly, the pooling layer in the network is replaced with a convolutional layer to reduce spatial loss of the fundus image. Then, we integrate multiscale feature fusion (MSFF) block into the encoders which helps the network to learn multiscale features efficiently and enrich the information carried with skip connection and lower-resolution decoder by fusing contextual channel attention (CCA) models. Finally, in order to solve the problems of data imbalance and misclassification, we present a Balanced Focal Loss function. In the experiments on benchmark dataset IDRID, we make an ablation study to verify the effectiveness of each component and compare FFU-Net against several state-of-the-art models. In comparison with baseline U-Net, FFU-Net improves the segmentation performance by 11.97%, 10.68%, and 5.79% on metrics SEN, IOU, and DICE, respectively. The quantitative and qualitative results demonstrate the superiority of our FFU-Net in the task of lesion segmentation of diabetic retinopathy.  
Web of Science 核心合集中的 "被引频次":  
36  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
7  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
37  
ISSN:  
2314-6133  
eISSN:  
2314-6141  
入藏号:  
WOS:000611824700002  
PubMed ID:  
33490274  
Indexed Date:  
2021-02-16  
Institution Address:  
Xi An Jiao Tong Univ, Sch Software, Xian 710054, Shaanxi, Peoples R China  
Institution Address:  
Huiyichen Inc, 1703 Block 1,1388 Jiulonghu Ave, Nanchang 330038, Jiangxi, Peoples R China  
Institution Address:  
Baidu Inc, Baidu Bldg,10 Shangdi 10th St, Beijing 100080, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, State Key Lab Mfg Syst Engn, Xian 710054, Peoples R China  
Institution:  
Institution:  
Huiyichen Inc  
Institution:  
Institution:

### 482. A method of single-shot target detection with multi-scale feature fusion

摘要: The Single Shot MultiBox Detector (SSD) is one of the fastest detection algorithms. Although it has achieved good results in detection, it also has the problem of poor detection effect for small targets and occlusion between objects. Here, the authors propose a new target detection method called single-shot target detection with multi-scale feature fusion and feature enhancement. Here, the authors introduce multi-scale feature fusion module, feature enhancement module and efficient channel attention module, and integrate them into the detection module of the original SSD target detection algorithm to improve the ability of network feature extraction. Experimental results on pascal VOC 2007 datasets show that the proposed algorithm works well when the input size is 300 x 300, the detection speed reaches 41.7 frames per second (FPS) and the detection accuracy reaches 79.6%, which is 2.4% higher than the original SSD target detection algorithm. When the input size is 512 x 512, the detection accuracy is 81.9%, and the detection speed reaches 36.5 FPS, which is 3.2% higher than the original SSD target detection algorithm. According to the experimental results, our algorithm has a better performance when there are many objects in the image and there is occlusion.  
Web of Science 核心合集中的 "被引频次":  
9  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
ISSN:  
1751-9659  
eISSN:  
1751-9667  
入藏号:  
WOS:000754468800001  
Indexed Date:  
2022-03-02  
Institution Address:  
Chongqing Univ Posts & Telecommun, Coll Comp Sci & Technol, 2 Chongwen Rd, Chongqing, Peoples R China  
Institution Address:  
Southwest Comp Co Ltd, Inst Informat Technol, Chongqing, Peoples R China  
Institution Address:  
Chongqing Coll Mobile Commun, 1 Vacat Rd, Chongqing, Peoples R China  
Institution:  
Institution:  
Southwest Comp Co Ltd  
Institution:  
Chongqing Coll Mobile Commun

### 483. PANet: A Point-Attention Based Multi-Scale Feature Fusion Network for

摘要: Point cloud registration is a critical task in many 3-D computer vision studies, aiming to find a rigid transformation that aligns one point cloud with another. In this article, we propose a point-attention based multi-scale feature fusion network (PANet) for partially overlapping point cloud registration. This study aims to investigate whether multi-scale features are more effective in improving the precision of alignment compared with fixed-scale local features. PANet comprises two core components: a multi-branch feature extraction module that extracts local features at different scales in parallel and a point-attention module (PAM) that learns an appropriate weight for each branch and then fuse these multi-scale features by weighted combination to enhance the representation ability of features. At the end of the network, four hidden layers are used to obtain the rigid transformation from the source point cloud to the template point cloud. Experiments on the synthetic ModelNet40 dataset demonstrate that the PANet outperforms state-of-the-art performance in terms of both alignment precision and robustness against noise. PANet also exhibits strong generalization ability on real-world Stanford 3-D and ICL-NUIM datasets. In addition, the computational complexity of our model compared to previous works is also evaluated. The results and ablation studies demonstrate that multi-scale fused local features are better at improving registration accuracy than fixed-scale local features. The findings may inspire future research in related fields and contribute to the development of new ideas and approaches.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
45  
在 BIOSIS Citation Index 中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
46  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000991806800007  
Indexed Date:  
2023-06-28  
Institution Address:  
Xidian Univ, Sch Comp Sci & Technol, Key Lab Collaborat Intelligence Syst, Minist Educ China, Xian 710071, Peoples R China  
Institution Address:  
Xidian Univ, Sch Elect Engn, Key Lab Collaborat Intelligence Syst China, Minist Educ China, Xian 710071, Peoples R China  
Institution Address:  
Xidian Univ, Sch Artificial Intelligence, Key Lab Intelligent Percept & Image Understanding, Minist Educ, Xian 710071, Peoples R China  
Institution:  
Institution:  
Institution:

### 484. A train bearing imbalanced fault diagnosis method based on extended CCR

摘要: The number of fault samples is much less than the normal samples in the actual operation of the train bearing, and the imbalanced characteristics of the fault data significantly decrease the performance of the diagnosis model. Therefore, a train bearing imbalanced fault diagnosis method (ECCR-MFFN) based on extended combined cleaning and resampling (ECCR) and the multi-scale feature fusion network (MFFN) is proposed. Firstly, the ECCR method is proposed, which adaptively determines the sampling area and provides rich fault information for the diagnostic model with high-quality synthesized samples. Then, MFFN is designed to obtain great feature extraction and classification results under imbalanced data conditions through feature extraction and fusion strategies of multi-branch different kernels. Finally, the superiority and effectiveness of the ECCR-MFFN under various data imbalance conditions are verified by comparative experiments on laboratory and public bearing datasets. The results demonstrate that the MFFN can effectively extract fault features under small imbalance rate (IBR) conditions and achieve ideal classification results. Compared with other data augmentation methods, the ECCR can synthesize samples with higher quality and has a more stable performance. Under the condition of IBR = 40:1, the accuracy of the ECCR-MFFN is 95.84% and 96.07%, which is significantly better than the comparison methods and offers a reliable method for dealing with data imbalance.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0924-090X  
eISSN:  
1573-269X  
入藏号:  
WOS:001231028000002  
Indexed Date:  
2024-05-31  
Institution Address:  
Guangxi Univ, Sch Mech Engn, Guangxi Key Lab Mfg Syst & Adv Mfg Technol, Nanning 530004, Peoples R China  
Institution Address:  
Jiaotong Univ, Sch Informat Sci & Engn Chongqing, Chongqing 400074, Peoples R China  
Institution Address:  
CRRC Zhuzhou Inst Co Ltd, Zhuzhou 412001, Peoples R China  
Institution:  
Institution:  
Jiaotong Univ  
Institution:

### 485. FBS-YOLO: an improved lightweight bearing defect detection algorithm

摘要: Aiming at the existing bearing defect detection algorithms with low accuracy, large number of parameters and computation, this paper proposes an efficient and lightweight bearing surface defect detection algorithm FBS-YOLO based on YOLOv8. Firstly, FasterNet replaces the original feature extraction network of YOLOv8, and uses Partial Convolution (PConv) to reduce redundant computation and memory access. Secondly, the fusion of weighted Bidirectional Feature Pyramid Network (BiFPN) in Neck network, which removes less efficient feature transmission nodes in the process of multi-scale feature fusion to achieve a higher level of fusion, improves the fusion efficiency of features at different scales. Finally, the advantages of Switchable Atrous Convolution (SAConv) are introduced to innovate the CSP Bottleneck with the two convolutions (C2f) module in the original model Neck network, and SAConv is combined with C2f (C2f\_SAConv) to from a more flexible module adapted to the features of different scales is proposed to enhance the feature extraction and processing capability of the model. The experimental results show that the algorithm FBS-YOLO proposed in this paper achieves a mAP of 91.4% in the bearing defect detection task, which is 2.8% higher than that of the original YOLO8 model, and the number of parameters and computation volume are reduced by 39.8% and 41.9%, respectively, and the model inference speed can be up to 161 fps. The algorithm meets the light-weight requirements of industrial detection deployment while maintaining high accuracy, effectively achieving a balance between model lightweight and performance, and providing new ideas for end-to-end industrial deployment.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0031-8949  
eISSN:  
1402-4896  
入藏号:  
WOS:001399201800001  
Indexed Date:  
2025-01-24  
Institution Address:  
Guilin Univ Technol, Educ Dept Guangxi Zhuang Autonomous Reg, Key Lab Adv Mfg & Automat Technol, Guilin 541006, Peoples R China  
Institution Address:  
Guilin Univ Technol, Coll Mech & Control Engn, Guilin 541006, Peoples R China  
Institution:  
Institution:

### 486. CMP-UNet: A Retinal Vessel Segmentation Network Based on Multi-Scale

摘要: Retinal vessel segmentation plays a critical role in the diagnosis and treatment of various ophthalmic diseases. However, due to poor image contrast, intricate vascular structures, and limited datasets, retinal vessel segmentation remains a long-term challenge. In this paper, based on an encoder-decoder framework, a novel retinal vessel segmentation model called CMP-UNet is proposed. Firstly, the Coarse and Fine Feature Aggregation module decouples and aggregates coarse and fine vessel features using two parallel branches, thus enhancing the model's ability to extract features for vessels of various sizes. Then, the Multi-Scale Channel Adaptive Fusion module is embedded in the decoder to realize the efficient fusion of cascade features by mining the multi-scale context information from these features. Finally, to obtain more discriminative vascular features and enhance the connectivity of vascular structures, the Pyramid Feature Fusion module is proposed to effectively utilize the complementary information of multi-level features. To validate the effectiveness of the proposed model, it is evaluated on three publicly available retinal vessel segmentation datasets: CHASE\_DB1, DRIVE, and STARE. The proposed model, CMP-UNet, reaches F1-scores of 82.84%, 82.55%, and 84.14% on these three datasets, with improvements of 0.76%, 0.31%, and 1.49%, respectively, compared with the baseline. The results show that the proposed model achieves higher segmentation accuracy and more robust generalization capability than state-of-the-art methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2079-9292  
入藏号:  
WOS:001116305900001  
Indexed Date:  
2023-12-16  
Institution Address:  
Henan Polytech Univ, Sch Comp Sci & Technol, Jiaozuo 454003, Peoples R China  
Institution Address:  
Southeast Univ, ST Yau Ctr, Sch Math, Nanjing 210096, Peoples R China  
Institution Address:  
Nanjing Ctr Appl Math, Nanjing 211135, Peoples R China  
Institution:  
Institution:  
Institution:

### 487. A novel network for semantic segmentation of landslide areas in remote

摘要: Landslides pose significant risks as natural disasters, highlighting the importance of accurate mapping using remote sensing images for various practical applications. However, due to the challenges arising from incomplete and inaccurate boundary information of foreground landslide polygons, existing methods can only achieve suboptimal performance. To this premise, in this paper, we propose a segmentation network called GMNet that leverages global information extraction and multi -scale feature fusion to enhance the discrimination of landslides from other objects. Specifically, by employing a multi -branch mechanism, our method effectively captures global information, while an improved multi -scale feature fusion technique addresses the issue of varying scales in landslide polygons. Furthermore, semantic enhancement enhances the semantic information of low-level features, bridging the semantic gap and enhancing fusion efficacy. Experimental results demonstrate the effectiveness of our network in segmenting landslide areas accurately within the remote sensing image dataset. Especially, our F1\_scores on three benchmarks outperform existing runner-ups by notable margins of 4.81%, 1.72%, and 1.16%, showcasing the value of our method in this domain.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1568-4946  
eISSN:  
1872-9681  
入藏号:  
WOS:001228893600001  
Indexed Date:  
2024-05-27  
Institution Address:  
Cent South Univ, Sch Automat, Changsha 410083, Peoples R China  
Institution Address:  
Great Bay Univ, Dongguan, Peoples R China  
Institution Address:  
Hunan Vocat Coll Engn, Dept Informat Engn, Changsha 410151, Peoples R China  
Institution:  
Institution:  
Institution:  
Hunan Vocat Coll Engn

### 488. Intelligent monitoring of small target detection using YOLOv8

摘要: In complex scenes, small target face detection is crucial but often hampered by detection accuracy and efficiency limitations. Our method addresses these challenges by incorporating Gaussian noise, which is key in improving model robustness and generalization. By simulating real-world imperfections, Gaussian noise acts as a regularizer and makes the model more resistant to variations in lighting and texture. Traditional methods often face difficulties when dealing with small targets and complex backgrounds due to inadequate feature extraction, suboptimal loss function design, and vulnerability to noise. To overcome these issues, we propose an improved YOLOv8 model based on multi-scale feature fusion and an optimized loss function. By leveraging Gaussian noise during training, our approach enhances both detection accuracy and operating efficiency. Experiments on the FDDB and WIDER FACE datasets demonstrate that our method performs better in various complex scenarios. Our method achieved 0.780 on the WIDER FACE validation set, outperforming existing mainstream techniques.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1110-0168  
eISSN:  
2090-2670  
入藏号:  
WOS:001364502000001  
Indexed Date:  
2024-12-03  
Institution Address:  
Suqian Univ, Sch Informat Engn, Suqian 223800, Peoples R China  
Institution Address:  
Suqian Univ, Ind Technol Res Inst, Suqian 223800, Peoples R China  
Institution:  
Institution:

### 489. Research on Waste Plastics Classification Method Based on Multi-Scale

摘要: Microplastic particles produced by non-degradable waste plastic bottles have a critical impact on the environment. Reasonable recycling is a premise that protects the environment and improves economic benefits. In this paper, a multi-scale feature fusion method for RGB and hyperspectral images based on Segmenting Objects by Locations (RHFF-SOLOv1) is proposed, which uses multi-sensor fusion technology to improve the accuracy of identifying transparent polyethylene terephthalate (PET) bottles, blue PET bottles, and transparent polypropylene (PP) bottles on a black conveyor belt. A line-scan camera and near-infrared (NIR) hyperspectral camera covering the spectral range from 935.9 nm to 1722.5 nm are used to obtain RGB and hyperspectral images synchronously. Moreover, we propose a hyperspectral feature band selection method that effectively reduces the dimensionality and selects the bands from 1087.6 nm to 1285.1 nm as the features of the hyperspectral image. The results show that the proposed fusion method improves the accuracy of plastic bottle classification compared with the SOLOv1 method, and the overall accuracy is 95.55%. Finally, compared with other space-spectral fusion methods, RHFF-SOLOv1 is superior to most of them and achieves the best (97.5%) accuracy in blue bottle classification.  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
eISSN:  
1424-8220  
入藏号:  
WOS:000873654300001  
PubMed ID:  
36298325  
Indexed Date:  
2022-11-04  
Institution Address:  
Fujian Prov Univ, Huaqiao Univ, Key Lab Proc Monitoring & Syst Optimizat Mech & E, Xiamen 361021, Peoples R China  
Institution:

### 490. Face Image Inpainting Method Based on Multi-Scale Feature Fusion

摘要: When dealing with severely corrupted images,the traditional image inpainting methods often produce blurry or too smooth regions in the restored images,and have difficulty in reconstructing a reasonable face image structure.To address the problem,this paper introduces a multi-scale feature fusion method into the discriminator of the traditional Generative Adversarial Network(GAN),which directly adds the upsampled feature maps of different depth to achieve effective fusion of shallow and deep information.Then this method grasps the overall pattern of the image with the help of high-level features,and fills the detail texture of the face image with low-level features,so that the resolution of the image can be fused with its semantic features for effective face image restoration.The experimental results on the dataset of CelebA show that the proposed method outperforms the regional normalization method in terms of Peak Signal to Noise Ratio(PSNR),structural similarity,L1 loss indicators,achieving ideal visual effects.  
摘要:  
传统图像修复方法在修复受损区域较大的图像时会出现修复结果过于平滑或模糊的现象,并且较难重建合理的人脸图像结构。在传统生成对抗网络的鉴别器中引入多尺度特征融合方法,将不同深度的特征图经过上采样后直接相加,使浅层信息和深层信息有效结合。通过借助高层特征把握图像的整体规律,同时利用低层特征填充人脸图像的细节纹理,进而使一张图像的分辨率及其语义特征相互融合,实现有效的人脸图像修复。在CelebA数据集上的实验结果表明,该方法的峰值信噪比、相似性结构、L1损失指标均优于区域归一化方法,取得了较好的视觉效果。  
在中国科学引文数据库中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
1000-3428  
入藏号:  
CSCD:6966584  
Indexed Date:  
2021-08-06  
Institution Address:  
延安大学物理与电子信息学院, 延安, 陕西 716000, 中国  
Institution Address:  
西安理工大学计算机学院, 西安, 陕西 710048, 中国  
Institution Address:  
School of Physics and Electronic Information,Yan'an University, Yan'an, Shaanxi 716000, China  
Institution Address:  
School of Computer Science and Engineering,Xi'an University of Technology, Xi'an, Shaanxi 710048, China  
Institution:  
延安大学物理与电子信息学院  
Institution:  
西安理工大学计算机学院  
Institution:  
School of Physics and Electronic Information,Yan'an University  
Institution:  
School of Computer Science and Engineering,Xi'an University of Technology

### 491. AMFF-LWBENet: A Novel Deep Learning Network Model for Extracting Lake

摘要: The extraction of lake water bodies from remote sensing images is vital for water resource management, environmental protection, climate change research, disaster monitoring, and land use planning. This study presents a novel model, AMFF-LWBENet (A Lake Water Body Extraction Network based on ASPP and Multi-scale Feature Fusion), designed to tackle challenges such as insufficient spatial detail, poor edge recognition, and low anti-noise performance found in existing lake water body extraction models. The model utilizes an encoder-decoder architecture, leveraging ResNet50 for downsampling to extract deep features, which are then fused more effectively through the multi-scale dense fusion (MDF) module, incorporating depthwise separable convolution and ASPP to enhance the integration of multi-scale features. Additionally, position and channel correlations in the feature map are captured using DANet, minimizing noise interference at the lake edge and improving segmentation accuracy. Bilinear interpolation is employed to upsample the feature map, and feature fusion is achieved through the cross-layer feature fusion (CFF) module, enabling precise lake water body extraction. The datasets utilized include a self-built W-H dataset, derived from Landsat images collected between 2015 and 2023, as well as the publicly available TP dataset. The AMFF-LWBENet achieved MIoU values of 97.52% on the W-H dataset and 97.5% on the TP dataset, surpassing current state-of-the-art semantic segmentation networks. These results suggest that AMFF-LWBENet is even more effective than many other networks in extracting lake water bodies, offering significantly enhanced support for a wide range of environmental and urban planning applications.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001337890000001  
Indexed Date:  
2024-10-26  
Institution Address:  
Inner Mongolia Agr Univ, Coll Comp & Informat Engn, Hohhot 010011, Peoples R China  
Institution Address:  
Inner Mongolia Autonomous Reg Key Lab Big Data Res, Hohhot 010011, Peoples R China  
Institution Address:  
Inner Mongolia Agr Univ, Coll Water Conservancy & Civil Engn, Hohhot 010018, Peoples R China  
Institution:  
Institution:  
Inner Mongolia Autonomous Reg Key Lab Big Data Res  
Institution:

### 492. Image Super-Resolution Reconstruction Based on Dense Residual Attention

摘要: In order to obtain super-resolution images with richer details and clearer textures, a method for image super-resolution reconstruction using dense residual attention and multi-scale fusion is proposed. First, different scale convolutions are used to fully extract shallow features of the image; then high-frequency features of the image are extracted through one three-layer cascaded multi-scale feature fusion and dense residual attention module, and the reuse of feature map is achieved; finally, residual branches are used to introduce shallow features and high-frequency features of each channel image, and the high-resolution images are reconstructed through up-sampling and sub-pixel convolution. The test results on the Set5, Set14, Bsd100, and Urban100 datasets show that the PSNR and SSIM of our model are superior to most current algorithms, especially in the case of x4 reconstruction results. PSNR has improved by 0.2 dB on the Set5 and Bsd100 datasets, and the algorithm has a better subjective visual effect.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0218-0014  
eISSN:  
1793-6381  
入藏号:  
WOS:001331698100001  
Indexed Date:  
2024-10-19  
Institution Address:  
Anhui Polytech Univ, Sch Comp & Informat, Wuhu 241000, Anhui, Peoples R China  
Institution:

### 493. Fault Diagnosis Method of Special Vehicle Bearing Based on Multi-Scale

摘要: To address the issues of inadequate feature extraction for rolling bearings, inaccurate fault diagnosis, and overfitting in complex operating conditions, this paper proposes a rolling bearing diagnosis method based on multi-scale feature fusion and transfer adversarial learning. Firstly, a multi-scale convolutional fusion layer is designed to effectively extract fault features from the original vibration signals at multiple time scales. Through a feature encoding fusion module based on the multi-head attention mechanism, feature fusion extraction is performed, which can model long-distance contextual information and significantly improve diagnostic accuracy and anti-noise capability. Secondly, based on the domain adaptation (DA) cross-domain feature adversarial learning strategy of transfer learning methods, the extraction of optimal domain-invariant features is achieved by reducing the gap in data distribution between the target domain and the source domain, addressing the call for research on fault diagnosis across operating conditions, equipment, and virtual-real migrations. Finally, experiments were conducted to verify and optimize the effectiveness of the feature extraction and fusion network. A public bearing dataset was used as the source domain data, and special vehicle bearing data were selected as the target domain data for comparative experiments on the effect of network transfer learning. The experimental results demonstrate that the proposed method exhibits an exceptional performance in cross-domain and variable load environments. In multiple bearing cross-domain transfer learning tasks, the method achieves an average migration fault diagnosis accuracy rate of up to 98.65%. When compared with existing methods, the proposed method significantly enhances the ability of data feature extraction, thereby achieving a more robust diagnostic performance.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001305742600001  
PubMed ID:  
39204877  
Indexed Date:  
2024-09-15  
Institution Address:  
Beijing Inst Technol, Sch Comp Sci & Technol, Beijing 100811, Peoples R China  
Institution Address:  
Changchun Univ, Coll Comp Sci & Technol, Changchun 130022, Peoples R China  
Institution Address:  
Natl Key Lab Special Vehicle Design & Mfg Integrat, Baotou 014030, Peoples R China  
Institution:  
Institution:  
Institution:  
Natl Key Lab Special Vehicle Design & Mfg Integrat

### 494. FFLKCDNet: First Fusion Large-Kernel Change Detection Network for

摘要: Change detection is an important technique that identifies areas of change by comparing images of the same location taken at different times, and it is widely used in urban expansion monitoring, resource exploration, land use detection, and post-disaster monitoring. However, existing change detection methods often struggle with balancing the extraction of fine-grained spatial details and effective semantic information integration, particularly for high-resolution remote sensing imagery. This paper proposes a high-resolution remote sensing image change detection model called FFLKCDNet (First Fusion Large-Kernel Change Detection Network) to solve this issue. FFLKCDNet features a Bi-temporal Feature Fusion Module (BFFM) to fuse remote sensing features from different temporal scales, and an improved ResNet network (RAResNet) that combines large-kernel convolution and multi-attention mechanisms to enhance feature extraction. The model also includes a Contextual Dual-Land-Cover Attention Fusion Module (CD-LKAFM) to integrate multi-scale information during the feature recovery stage, improving the resolution of details and the integration of semantic information. Experimental results showed that FFLKCDNet outperformed existing methods on datasets such as GVLM, SYSU, and LEVIR, achieving superior performance in metrics such as Kappa coefficient, mIoU, MPA, and F1 score. The model achieves high-precision change detection for remote sensing images through multi-scale feature fusion, noise suppression, and fine-grained information capture. These advancements pave the way for more precise and reliable applications in urban planning, environmental monitoring, and disaster management.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001442488800001  
Indexed Date:  
2025-03-18  
Institution Address:  
Macao Polytech Univ, Fac Appl Sci, Macau 999078, Peoples R China  
Institution Address:  
Macao Polytech Univ, Macau 999078, Peoples R China  
Institution:  
Institution:

### 495. Acoustic tomography temperature reconstruction based on improved sparse

摘要: Acoustic tomography is a widely used non -contact method for visualizing temperature distribution. A temperature distribution reconstruction algorithm based on an improved sparse reconstruction model and multi -scale feature fusion network is proposed. First, the acoustic temperature measurement sparse reconstruction model is improved by combining the error function (ERF) and the iterative reweighting algorithm, and the alternating direction method of multipliers algorithm (ADMM) is used to solve the model to obtain the initial temperature distribution. Then the feature extraction network is constructed to extract multi -scale features of acoustic time of flight (TOF) as prior information. Finally, the feature fusion reconstruction network is constructed to fuse and reconstruct the initial temperature distribution and multi -scale features to obtain a high -precision temperature distribution. Simulation and experimental tests were conducted respectively, and compared with other algorithms. The results show that the average relative error and root mean square error of the simulated temperature distribution reconstruction are 0.073% and 0.1% respectively, the average reconstruction error of the temperature points set in the experimental test is 0.38%, and the reconstruction errors are lower than other algorithms. The proposed method effectively utilizes prior information to correct sparse reconstruction temperature distribution reconstruction results, significantly improving the quality of temperature distribution reconstruction.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001209815600001  
Indexed Date:  
2024-05-04  
Institution Address:  
North China Elect Power Univ, Dept Automat, Baoding 071003, Peoples R China  
Institution Address:  
Baoding Key Lab State Detect & Optimizat Regulat I, Baoding 071003, Peoples R China  
Institution:  
Institution:  
Baoding Key Lab State Detect & Optimizat Regulat I

### 496. Lightweight road extraction model based on multi-scale feature fusion

摘要: A road extraction model based on multi-scale feature fusion lightweight DeepLab V3+ (MFL-DeepLab V3+) was proposed aiming at the problems of high computational complexity and poor road extraction effect of the current semantic models used in the field of remote sensing image road extraction. The lightweight MobileNet V2 network was used to replace the original models Xception network as the backbone network in order to reduce the parameters of the model and the computational complexity of the model. Deep separable convolution was introduced into the Atlas spatial pyramid pooling (ASPP) module. A multi-scale feature fusion with attention (MFFA) was proposed in the decoding area in order to enhance the road extraction ability of the model and optimize the extraction effect on small road segments. Experiments based on the Massachusetts roads dataset showed that the parameter size of the MFL-DeepLab V3+ model was significantly reduced with a parameter compression of 88.67% compared to the original model. The road extraction image had clear edges, and its accuracy, recall, and F1-score were 88.45%, 86.41% and 87.42%, achieving better extraction performance compared to other models.  
摘要:  
针对当前用于遥感图像道路提取领域的语义模型存在计算复杂度较高、道路提取效果不佳的问题,提出基于多尺度特征融合的轻量化道路提取模型(MFL-DeepLab V3+).为了减少模型参数量并降低模型的计算复杂度,骨干网络选用轻量化Mobilenet V2网络代替原模型的Xception网络,在空洞空间金字塔池化(ASPP)模块中引入深度可分离卷积.为了增强模型的道路提取能力,优化对细小路段的提取效果,在解码区提出联合注意力的多尺度特征融合(MFFA).基于Massachusetts roads数据集的各项实验表明, MFL-DeepLab V3+模型的参数规模显著降低,较原模型参数量压缩了88.67%,道路提取图像完整,边缘清晰,精确率、召回率和F1分数分别达到88.45%、 86.41%和87.42%,与其他模型相比取得了更好的提取效果.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1008-973X  
入藏号:  
CSCD:7701297  
Indexed Date:  
2024-06-29  
Institution Address:  
天津城建大学计算机与信息工程学院, 天津 300384, 中国  
Institution Address:  
School of Computer and Information Engineering, Tianjin Chengjian University, Tianjin 300384, China  
Institution:  
天津城建大学计算机与信息工程学院  
Institution:  
School of Computer and Information Engineering, Tianjin Chengjian University

### 497. Improved Fabric Defect Detection Algorithm of YOLOv5

摘要: Aiming at the problems of slow detection speed of two-stage algorithm and low detection accuracy of one-stage algorithm in the current network model applied to fabric defect detection, an improved YOLOv5 fabric defect detection algorithm is proposed. Firstly, for the different sizes of fabric defects, the clustering distance standard of the [K]-mean algorithm is modified, and the size of the priori frame is recalculated. Secondly, the standard convolution(SC) of the network Neck layer is improved, and the depth separation convolution(DSC) is combined with the standard convolution to reduce the amount of network layer parameters while maintaining the feature extraction capability of the network. The coordinate attention(CA) mechanism is introduced in the feature fusion stage, so that the network can capture the connection between each channel while retaining the precise positioning information of the target, thereby enhancing the feature extraction and positioning capabilities of the network. Finally, the weighted bidirectional feature pyramid network(BiFPN) is used, the feature pyramid module is modified to achieve simple and fast multi-scale feature fusion. After training on the data set, the results show that the mAP value of the improved YOLOv5 model can reach 97.4%, which is 2.8 percentage points higher than the original network accuracy, which meets the requirements of fabric defect detection.  
摘要:  
针对目前应用到织物疵点检测的网络模型中存在的两阶段算法检测速度慢、一阶段算法检测精度低的问题,提出了一种改进YOLOv5的织物疵点检测算法。针对织物疵点大小不一的问题,对[K]-mean算法的聚类距离标准进行修改,重新计算先验框大小;对网络Neck层标准卷积(standard convolution,SC)进行改进,将深度分离卷积(depth separation convolution,DSC)与标准卷积结合,减少网络层参数量,同时保持网络的特征提取能力;在特征融合阶段引入坐标注意力机制(coordinate attention,CA),使网络能够捕捉各通道之间联系的同时保留目标的精确定位信息,加强网络的特征提取和定位能力;使用加权双向特征金字塔网络(bidirectional feature pyramid network,BiFPN)中的方法,对特征金字塔模块进行修改,实现简单快速的多尺度特征融合。在数据集上进行训练,结果表明,改进的YOLOv5模型的mAP值可达到97.4%,相比于原网络精度提高了2.8个百分点,满足了织物疵点检测的要求。  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1002-8331  
入藏号:  
CSCD:7494742  
Indexed Date:  
2023-09-28  
Institution Address:  
西安工程大学纺织科学与工程学院, 西安, 陕西 710048, 中国  
Institution Address:  
College of Textile Science and Engineering, Xi'an Polytechnic University, Xi'an, Shaanxi 710048, China  
Institution:  
西安工程大学纺织科学与工程学院  
Institution:  
College of Textile Science and Engineering, Xi'an Polytechnic University

### 498. Double-view feature fusion network for LiDAR semantic segmentation

摘要: Objective Point cloud semantic segmentation, as the basic technology of 3D point cloud data target detection, point cloud classification, and other projects, is an important part of the current 3D computer vision. At the same time, point cloud segmentation technology is the key for the computer to understand scenes, and it has been widely used in many fields such as autonomous driving, robotics, and augmented reality. Point cloud semantic segmentation refers to the pointby- point classification operation of points in the point cloud scene, that is, to judge the category of each point in the point cloud and finally segment and integrate accordingly. Generally, point cloud semantic segmentation technology can be divided into two categories according to different application scenarios: small-scale point cloud semantic segmentation and large-scale point cloud semantic segmentation. Small-scale point cloud semantic segmentation only performs semantic segmentation operations on indoor point cloud scenes or small-scale point cloud scenes, whereas the large-scale point cloud semantic segmentation replaces the deployment environment of the algorithm with outdoor large-scale point cloud data. Classification and integration for point clouds are usually performed on driving scenes or urban scenes. Compared with the point cloud semantic segmentation of small scenes, the semantic segmentation of large-scale point clouds has a wider range of applications and is extensively used in driving scene understanding, urban scene reconstruction, and other fields. However, due to the large amount of data and the complexity of point cloud data, the task of semantic segmentation for point cloud in large scenes is more difficult. To improve the extraction quality of point features in a large-scale point cloud, a semantic segmentation method based on double-view feature fusion network for LiDAR semantic segmentation is proposed. Method Our method is composed of two parts, double-view feature fusion module and feature integration based on asymmetric convolution. In the down sampling stage, a double-view feature fusion module, which includes a double-view point cloud feature extraction module and a feature fusion block, is suggested. The double-view feature fusion module combines the cylindrical feature with the global feature of key points to reduce the feature loss caused by downsampling. The features in different views of the point cloud are combined by feature splicing in this module. Finally, the combined point cloud features are placed into the feature fusion block for feature dimensionality reduction and fusion. In the feature integration stage, a point cloud feature integration module is proposed based on asymmetric convolution, including asymmetric convolution and multiscale dimension-decomposition context modeling, achieving the enhancement and reconstruction of point cloud features by the operation of asymmetric point cloud feature processing and multi-scale context feature integration. The feature integration processes the double-view feature by asymmetric convolution and then uses multi-dimensional convolution and multiscale feature integration for feature optimization. Result In our experimental environment, our algorithm has the second-highest frequency weighted intersection over union accuracy rate and the highest mean intersection over union (mIoU) accuracy rate among recent algorithms. Our work focuses on the improvement of segmentation accuracy and achieves the highest segmentation accuracy in multiple categories. In vehicle categories such as cars and trucks, our method achieves a high segmentation accuracy. In categories such as bicycles, motorcycles, and pedestrians with small individuals and complex shapes, our method performs better than other methods.  
摘要:  
目的点云语义分割在无人驾驶、城市场景建模等领域中具有重要意义,为了提升大场景条件下点云特征的提取效率,提出一种大场景双视角点云特征融合的语义分割方法(double-view feature fusion network for LiDAR semantic segmentation,DVFNet)。方法大场景双视角点云特征融合的语义分割方法由两个部分组成,分别为双视角点云特征融合模块和基于非对称卷积的点云特征整合模块。双视角点云特征融合模块将柱状体素特征与关键点全局特征相结合,减少降采样导致的特征损失;基于非对称卷积的点云特征整合模块将双视角点云特征使用非对称卷积进行处理,并使用多维度卷积与多尺度特征整合来实现局部特征优化。结果本文提出的大场景双视角点云特征融合语义分割方法,在SemanticKITTI大场景点云数据集上达到63.9%的准确率,分割精度在已开源的分割方法中处于领先地位。结论通过文中的双视角点云特征融合语义分割方法,能够实现大场景条件下点云数据的高精度语义分割。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1006-8961  
入藏号:  
CSCD:7664640  
Indexed Date:  
2024-05-17  
Institution Address:  
上海理工大学出版印刷与艺术设计学院, 上海 200093, 中国  
Institution Address:  
College of Communication and Art Design,University of Shanghai for Science and Technology, Shanghai 200093, China  
Institution:  
上海理工大学出版印刷与艺术设计学院  
Institution:  
College of Communication and Art Design,University of Shanghai for Science and Technology

### 499. Few-shot based learning recaptured image detection with multi-scale

摘要: Advancements in LCD screen resolution and digital camera technology have significantly improved the quality of recaptured images, making it increasingly challenging to differentiate them from original images. Previous approaches to recaptured image detection have primarily focused on generating large-scale datasets of recaptured images and utilizing deep-learning techniques for detection. However, in rare and emerging scenarios, such as detecting image manipulation in medical images associated with rare diseases, identifying forgeries in precious art, and addressing new social media trends, the availability of recaptured image datasets is often limited. To address these challenges, a few-shot learning approach for recaptured image detection is proposed, which effectively captures local and global information by introducing attention mechanisms at different scales. This enhances the feature representation capability and enables the identification of subtle differences between recaptured and original images. Three complementary modules are designed: parallel multi-scale feature fusion, cascade multi-scale feature fusion, and an attention mechanism, to improve feature representation, capture key information, and highlight significant features, thereby enhancing detection accuracy. Experimental results confirm the superior performance of our method in detecting recaptured images, particularly with limited training samples, effectively addressing the challenges in detecting recaptured images with few labeled samples.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:001385579800001  
Indexed Date:  
2025-01-03  
Institution Address:  
Shenzhen Univ, Coll Comp Sci & Software Engn, Shenzhen 518060, Peoples R China  
Institution Address:  
Shenzhen City Polytech, Sch Informat & Commun Technol, Shenzhen 518116, Peoples R China  
Institution Address:  
Shenzhen Univ, Shenzhen Key Lab Media Secur, Shenzhen 518060, Peoples R China  
Institution Address:  
Shenzhen MSU BIT Univ, Fac Engn, Guangdong Lab Machine Percept & Intelligent Comp, Shenzhen 518116, Peoples R China  
Institution:  
Institution:  
Shenzhen City Polytech  
Institution:  
Institution:

### 500. MFDC-Net: A breast cancer pathological image classification algorithm

摘要: Breast cancer is one of the most common malignant tumors in the world. Traditional methods take pathologists a lot of time and effort to diagnose, and the results are greatly affected by individual abilities. Using computer-aided diagnosis methods can improve the accuracy and efficiency of pathological image classification, meet the demands of clinical applications. To this end, a multi-scale feature fusion based on DenseNet and coordinate attention network(MFDC-Net)is proposed. The introduction of coordinate attention mechanism into the dense blocks can locate important feature spatial information precisely. The improved transition layers use average pooling and normal convolutions with different convolution kernels to reduce dimension and expand receptive fields. Finally the improved network employs a multi-scale feature fusion model using dilated convolution, average pooling and normal convolutions to fuse deep image features to improve classification performance. The experimental results show that MFDC-Net model has better classification performance, the accuracy rate of four classifications reaches 97.12%, the easily confused rate decreases to 3.34%. The method can better classify the histopathological images of breast cancer, and can provide an important basis for the diagnosis and treatment of doctors.  
摘要:  
乳腺癌是全球最常见的恶性肿瘤之一,采用传统方法诊断需花费大量时间和精力,且受个人能力影响较大。用计算机辅助诊断的方法,可以提高病理图像分类的准确率和效率,从而满足临床应用的需求。为此,提出一种基于DenseNet的融合多尺度特征和注意力机制的乳腺癌病理图像分类算法(MFDC-Net)。在密集块中引入坐标注意力机制,精准定位重要特征的空间信息。采用多尺度池化过渡层,通过不同卷积核的平均池化和普通卷积,在实现降维的同时扩大感受野。采用多尺度特征增强模块,融合深层次图像特征,提高分类性能。结果显示,MFDC-Net模型的分类性能较其他经典模型更优,分类准确率达97.12%,易混淆率低至3.34%,能较好地进行乳腺癌组织病理图像分类,为诊断和治疗提供重要依据。  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1008-9497  
入藏号:  
CSCD:7512889  
Indexed Date:  
2023-10-28  
Institution Address:  
浙江工业大学管理学院, 杭州, 浙江 310023, 中国  
Institution Address:  
School of Management, Zhejiang University of Technology, Hangzhou, Zhejiang 310023, China  
Institution:  
浙江工业大学管理学院  
Institution:  
School of Management, Zhejiang University of Technology

### 501. Aero-engine prognosis strategy based on multi-scale feature fusion and

摘要: Aero-engine prognosis is helpful to ensure its safety and reliability, and effectively reduce the maintenance cost. However, the existing works only perform RUL prediction, ignoring the fault factors that lead to engine degradation. In addition, most prognosis methods can only extract single-scale features, ignoring the potential degradation features at other scales and layers. Therefore, this work proposes an aero-engine prognosis framework based on multi-scale feature fusion and multi-task parallel learning. In the proposed framework, multi-scale feature fusion blocks are designed to explore and fuse the potential degradation features of samples under different scales. And a layers concatenation block is constructed to integrate feature details from different layers and avoid losing useful information. Then a multi-task parallel learning block is constructed, and a joint loss function is developed for parallel learning of RUL prediction and fault diagnosis tasks. Meanwhile, a stacked image conversion method is proposed to integrate multi-sensor data with multiple cycles into image sample and make it contains more information beneficial to engine degradation. Finally, experimental results on CMAPSS and N-CMAPSS datasets show that the proposed framework exhibits superiority over other state-of-the-art methods and demonstrates good generalization and robustness.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
24  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
24  
ISSN:  
0951-8320  
eISSN:  
1879-0836  
入藏号:  
WOS:000946733400001  
Indexed Date:  
2023-03-27  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Civil Aviat, Nanjing 211100, Peoples R China  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Civil Aviat, Nanjing 211100, Peoples R China  
Institution Address:  
Shandong Normal Univ, Sch Publ Adm, Jinan 250000, Peoples R China  
Institution:  
Institution:  
Institution:

### 502. Depth Image Super-Resolution Reconstruction Network Based on Dual

摘要: A depth image super-resolution reconstruction network (DF-Net) based on dual feature fusion guidance is proposed to address the issues of texture transfer and depth loss in color image guided deep image super-resolution reconstruction algorithms. To fully utilize the correlation between depth and intensity features, a dual channel fusion module (DCM) and a dual feature guided reconstruction module (DGM) are used to perform deep recovery and reconstruction in the network model. The multi-scale features of depth and intensity information are extracted using a input pyramid structure: DCM performs feature fusion and enhancement between channels based on a channel attention mechanism for depth and intensity features; DGM provides dual feature guidance for reconstruction by adaptively selecting and fusing depth and intensity features, increasing the guidance effect of depth features, and overcoming the issues of texture transfer and depth loss. The experimental results show that the peak signal-to-noise ratio (PSNR) and root mean square error (RMSE) of the proposed method are superior to those of methods such as RMRF, JBU, and Depth Net. Compared to the other methods, the PSNR value of the 4x super-resolution reconstruction results increased by an average of 6. 79 dB, and the RMSE decreased by an average of 0. 94, thus achieving good depth image super- resolution reconstruction results.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001283351300042  
Indexed Date:  
2024-09-19  
Institution Address:  
Changchun Univ Sci & Technol, Coll Informat Sci & Engn, Changchun 130012, Jilin, Peoples R China  
Institution:

### 503. Scene Uyghur Text Detection Based on Fine-Grained Feature Representation

摘要: Scene text detection task aims to precisely localize text in natural environments. At present, the application scenarios of text detection topics have gradually shifted from plain document text to more complex natural scenarios. Objects with similar texture and text morphology in the complex background noise of natural scene images are prone to false recall and difficult to detect multi-scale texts, a multi-directional scene Uyghur text detection model based on fine-grained feature representation and spatial feature fusion is proposed, and feature extraction and feature fusion are improved to enhance the network's ability to represent multi-scale features. In this method, the multiple groups of 3 x 3 convolutional feature groups that are connected like the hierarchical residual to build a residual network for feature extraction, which captures the feature details and increases the receptive field of the network to adapt to multi-scale text and long glued dimensional font detection and suppress false positives of text-like objects. Secondly, an adaptive multi-level feature map fusion strategy is adopted to overcome the inconsistency of information in multi-scale feature map fusion. The proposed model achieves 93.94% and 84.92% F-measure on the self-built Uyghur dataset and the ICDAR2015 dataset, respectively, which improves the accuracy of Uyghur text detection and suppresses false positives.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
eISSN:  
1424-8220  
入藏号:  
WOS:000816360500001  
PubMed ID:  
35746154  
Indexed Date:  
2022-07-03  
Institution Address:  
Xinjiang Univ, Sch Informat Sci & Engn, Urumqi 830046, Peoples R China  
Institution Address:  
Xinjiang Univ, Xinjiang Key Lab Multilingual Informat Technol, Urumqi 830046, Peoples R China  
Institution:  
Institution:

### 504. A novel mooring system anomaly detection framework for SEMI based on

摘要: The structural safety of mooring line is of paramount importance for maintaining the stability of floating structure and personnel health. Once mooring line failure occurs, it may lead to catastrophic consequences. Realtime monitoring and damage identification of mooring line integrity provide an early warning and response to mitigate potential risks and losses. This paper presents a motion-based mooring line anomaly detection framework, combining continuous wavelet transform, multi-scale feature fusion, and squeeze-and-excitation residual network (namely CWT-FFSeResNet). The framework aims to identify different degrees of mooring line damage in a semi-submersible platform (SEMI). Extensive numerical simulations under various sea conditions provide motion response data for different mooring line damage states. Subsequently, time-series motion data is converted into a time-frequency image, and feature fusion stacks images of three motions from the same time period on channel, forming a whole sample to represent the state of a mooring line. Compared with other existing models, the model shows a perfect performance in terms of accuracy and efficiency. Based on the test results of insufficient samples, the model indicates the potential to be established at a smaller time consuming. In addition, test experiments with different Gaussian noise levels demonstrated relatively satisfactory noise robustness of proposed method.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
7  
ISSN:  
0951-8320  
eISSN:  
1879-0836  
入藏号:  
WOS:001179573400001  
Indexed Date:  
2024-03-27  
Institution Address:  
China Univ Petr, Coll Safety & Ocean Engn, Beijing, Peoples R China  
Institution Address:  
Tsinghua Univ, Inst Ocean Engn, Shenzhen Int Grad Sch, Shenzhen, Peoples R China  
Institution Address:  
China Univ Petr, Coll Petr Engn, Beijing, Peoples R China  
Institution Address:  
Chinese Acad Sci, LHD, Inst Mech, Beijing 100190, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 505. AHC-Net: a road crack segmentation network based on dual attention

摘要: To solve the problem of incomplete and inaccurate pavement crack detection, an improved U-Net model based on dual attention mechanism and multi-feature fusion is proposed. Firstly, a new encoding module ACI is designed, which has the feature of multi-scale feature extraction, significantly improves the sensing ability of the damaged area, reduces the background interference, and realizes more accurate segmentation. Secondly, a new decoding module HAD is designed, which avoids the network degradation problem caused by gradient vanishing and the growth of network layers and can retain the most subtle feature information during the decoding process. Finally, convolutional block attention module (CBAM) is introduced in the encoding part to effectively extract global and local detail information, and the criss-cross attention mechanism is also introduced in the decoding part to prevent the loss of marginalized information. The model proposed in this article was tested on the public datasets DeepCrack, CrackSeg478, and AsphaltCrack300, and compared with other advanced methods. The experimental results indicate that this method can detect road cracks more accurately and possesses considerable robustness.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001222391000001  
Indexed Date:  
2024-05-23  
Institution Address:  
North China Univ Sci & Technol, Hebei Key Lab Ind Intelligent Percept, Tangshan 063210, Peoples R China  
Institution Address:  
Tangshan Univ, Intelligence & Informat Engn Coll, Tangshan 063000, Peoples R China  
Institution Address:  
Zhejiang A&F Univ, Coll Math & Comp Sci, Hangzhou 311300, Peoples R China  
Institution:  
Institution:  
Institution:

### 506. A sub-pixel convolution-based improved bidirectional feature pyramid

摘要: Due to the dramatic scale changes in different ground object characteristics, effectively utilizing multi-scale features of remote sensing images represents a major challenge. This letter proposes a sub-pixel convolution-based improved bidirectional feature pyramid network (SCIBFPN) to address the problem. To reduce the information loss in the process of image pre-processing and multi-scale feature fusion, this letter introduces sub-pixel convolution instead of up-sampling. To obtain richer spatial and spectral information, an improved bidirectional feature pyramid network-based sub-pixel convolution is combined with a residual network (ResNet) for feature extraction. In addition, to further enhance the spatial structure of the fused images, a panchromatic (PAN) image is used as a guide to direct the injection of spatial information. Experimental results from two real datasets show that the proposed method outperforms state-of-the-art methods in terms of both objective metrics and subjective visual evaluation.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2150-704X  
eISSN:  
2150-7058  
入藏号:  
WOS:000905892700001  
Indexed Date:  
2023-01-18  
Institution Address:  
Zhengzhou Univ Light Ind, Coll Comp & Commun Engn, Zhengzhou, Peoples R China  
Institution Address:  
Acad Forest Inventory & Planning, Natl Forestry & Grassland Adm, Beijing, Peoples R China  
Institution Address:  
Zhengzhou Univ Light Ind, Coll Comp & Commun Engn, Zhengzhou 450001, Peoples R China  
Institution:  
Institution:  
Acad Forest Inventory & Planning  
Institution:

### 507. 被撤回的出版物: A Pavement Crack Detection Method Based on Multiscale Attention

摘要: To solve the problem of low detection accuracy due to the loss of detailed information when extracting pavement crack features in traditional U-shaped networks, a pavement crack detection method based on multiscale attention and hesitant fuzzy set (HFS) is proposed. First, the encoding-decoding structure is used to construct a pavement crack segmentation network, ResNeXt50 is used to extract features in the encoding stage, and a multiscale feature fusion module (MFF) is designed to obtain multiscale context information. Second, in the decoding stage, a high-efficiency dual attention module (EDA) is used to enhance the ability of capturing details of the cracks while suppressing background noise. Finally, the membership degree of the crack is calculated based on the advantages of the HFS in multiattribute decision-making to obtain the similarity of the crack, and the binary image after segmentation is judged by the hesitation fuzzy measure. The experiment was conducted on the public road crack dataset Crack500. In terms of segmentation performance, the evaluation indexes Intersection over Union (IoU), Precision, and Dice coefficients of the proposed network reached 55.56%, 74.26%, and 67.43%, respectively; in terms of classification performance, for transversal and longitudinal cracks, the classification accuracy was 84%+/- 0.5%, while the block and the alligator were both 78%+/- 0.5%. The experimental results prove that the crack details detected by the proposed method are more abundant, and the image detection effect of complex topological structures and small cracks are better.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
1687-5265  
eISSN:  
1687-5273  
入藏号:  
WOS:000783317700007  
PubMed ID:  
35126484  
Indexed Date:  
2022-04-24  
Institution Address:  
Hebei Univ, Sch Cyber Secur & Comp, Baoding 071002, Peoples R China  
Institution Address:  
Hebei Univ, Hebei Machine Vis Engn Res Ctr, Baoding 071002, Peoples R China  
Institution Address:  
Hebei Univ, Inst Intelligent Image & Document Informat Proc, Baoding 071002, Peoples R China  
Institution Address:  
Hebei Univ, Sch Math & Informat Sci, Baoding 071002, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 508. Trident-LK Net: A Lightweight Trident Structure Network With Large

摘要: Identifying defects at different scales is a challenge in industrial defect detection. To solve this problem, many multi-scale feature fusion networks have been proposed to improve multi-scale target detection accuracy by fusing fine-grained information from shallow networks and semantic information from deep networks. This approach requires the introduction of extraa parameters. Thinking from another perspective, can the accuracy of multi-scale target detection be improved by fusing the feature information under different receptive fields? For this purpose, we designed a three-layer network structure called Trident-LK Net. our model uses convolutional kernels of different sizes (31, 25, 1) in the feature extraction phase and establishes cross-fusion connections. This omits the feature fusion part and greatly reduces the network parameters while obtaining a good detection accuracy. Finally we perform experiments on the neu-det dataset and the gc10 dataset to verify the feasibility of our idea. While keeping the number of parameters to a minimum, our model achieves competitive detection results on the neu-det dataset (76.9% mAP) and optimal on the gc10 dataset (63.55% mAP). Our code will be publicly available at https://github.com/syyang2022/Trident-LK-Net.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
2169-3536  
入藏号:  
WOS:001121204200001  
Indexed Date:  
2023-12-23  
Institution Address:  
Lanzhou Univ Technol, Sch Mech & Elect Engn, Lanzhou 730050, Peoples R China  
Institution Address:  
Zhejiang Univ Sci & Technol, Hangzhou 310023, Zhejiang, Peoples R China  
Institution Address:  
Zhejiang Dingli Ind Co Ltd, Lishui 321400, Zhejiang, Peoples R China  
Institution:  
Institution:  
Institution:  
Zhejiang Dingli Ind Co Ltd

### 509. LHFFNet: A hybrid feature fusion method for lane detection

摘要: Lane line images have the essential attribute of large-scale variation and complex scene information, and the similarity between adjacent lane lines is high, which can easily cause classification errors. And remote lane lines are difficult to recognize due to visual angle changes in width. To address this issue, this paper proposes an effective lane detection framework, which is a hybrid feature fusion network that enhances multiple spatial features and distinguishes key features throughout the entire lane line segment. It enhances and fuses lane line features at multiscale to enhance the feature representation of lane line images, especially at the far end. Firstly, in order to enhance the correlation of multiscale lane features, a multi-head self attention is used to construct a multi-space attention enhancement module for feature enhancement in multispace. Secondly, a spatial separable convolutional branch is designed for the jumping layer structure connecting multiscale lane line features. While retaining feature information of different scales, important lane areas in multiscale feature information are emphasized through the allocation of spatial attention weights. Finally, considering that lane lines are elongated areas in the image, and the background information in the image is much more abundant than lane line information, the flexibility of traditional pooling operations in capturing widely existing anisotropic contexts in actual environments is limited. Therefore, before embedding feature output branches, strip pooling is introduced to refine the representation of lane line information and optimize model performance. The experimental results show that the accuracy on the TuSimple dataset reaches 96.84%, and the F1 score on the CULane dataset reaches 75.9%.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2045-2322  
入藏号:  
WOS:001337302400069  
PubMed ID:  
39013975  
Indexed Date:  
2025-01-17  
Institution Address:  
Cent South Univ Forestry & Technol, Sch Comp & Math, Changsha 410004, Hunan, Peoples R China  
Institution Address:  
Cent South Univ Forestry & Technol, Sch Elect Informat & Phys, Changsha 410004, Hunan, Peoples R China  
Institution:  
Institution:

### 510. Enhancing Underwater Video from Consecutive Frames While Preserving

摘要: Current methods for underwater image enhancement primarily focus on single-frame processing. While these approaches achieve impressive results for static images, they often fail to maintain temporal coherence across frames in underwater videos, which leads to temporal artifacts and frame flickering. Furthermore, existing enhancement methods struggle to accurately capture features in underwater scenes. This makes it difficult to handle challenges such as uneven lighting and edge blurring in complex underwater environments. To address these issues, this paper presents a dual-branch underwater video enhancement network. The network synthesizes short-range video sequences by learning and inferring optical flow from individual frames. It effectively enhances temporal consistency across video frames through predicted optical flow information, thereby mitigating temporal instability within frame sequences. In addition, to address the limitations of traditional U-Net models in handling complex multiscale feature fusion, this study proposes a novel underwater feature fusion module. By applying both max pooling and average pooling, this module separately extracts local and global features. It utilizes an attention mechanism to adaptively adjust the weights of different regions in the feature map, thereby effectively enhancing key regions within underwater video frames. Experimental results indicate that when compared with the existing underwater image enhancement baseline method and the consistency enhancement baseline method, the proposed model improves the consistency index by 30% and shows a marginal decrease of only 0.6% in enhancement quality index, demonstrating its superiority in underwater video enhancement tasks.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2077-1312  
入藏号:  
WOS:001404235200001  
Indexed Date:  
2025-02-01  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Automat, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Jiangsu Collaborat Innovat Ctr Atmospher Environm, Nanjing 210044, Peoples R China  
Institution Address:  
Univ Reading, Whiteknights, POB 217, Reading RG6 6AH, England  
Institution Address:  
State Grid Jiangsu Elect Power Co, Informat & Telecommun Branch, Nanjing 211125, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 511. Double-branch forgery image detection based on multi-scale feature

摘要: Most of existing methods exhibit poor performance in detecting forged images due to the small size of tampered areas and the limited pixel difference between untampered and tampered regions. To alleviate the above problem, a double-branch tampered image detection based on multi-scale features is proposed. Firstly, we introduce a fusion module based on attention mechanism in the first branch to enhance the network's sensitivity towards tampered regions. Secondly, we construct a second branch specifically designed for detection, aiming to identify subtle differences between tampered and untampered areas by utilizing rich edge information from shallow features as guidance. Compared to the existing methods on the public benchmark datasets CASIA1.0, Columbia and NIST16, the values of F-score reached 0.766, 0.900 and 0.930 on those datasets, respectively. The experimental results show that our method could significantly improve the accuracy on detecting the tampered area.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1673-1905  
eISSN:  
1993-5013  
入藏号:  
WOS:001195640200008  
Indexed Date:  
2024-04-12  
Institution Address:  
Civil Aviat Univ China, Coll Elect Informat & Automat, Tianjin 300300, Peoples R China  
Institution:

### 512. BEMF-Net: Semantic Segmentation of Large-Scale Point Clouds via

摘要: The semantic segmentation of point clouds is a crucial undertaking in 3D reconstruction and holds great importance. However, achieving precise semantic segmentation represents a significant hurdle. In this paper, we present BEMF-Net, an efficient method for large-scale environments. It starts with an effective feature extraction method. Unlike images, 3D data comprise not only geometric relations but also texture information. To accurately depict the scene, it is crucial to take into account the impacts of texture and geometry on the task, and incorporate modifications to improve feature description. Additionally, we present a multi-scale feature fusion technique that effectively promotes the interaction between features at different resolutions. The approach mitigates the problem of the smoothing of detailed information caused by downsampling mechanisms, while ensuring the integrity of features across different layers, allowing a more comprehensive representation of the point cloud. We confirmed the effectiveness of this method by testing it on benchmark datasets such as S3DIS, SensatUrban, and Toronto3D.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001113719500001  
Indexed Date:  
2023-12-17  
Institution Address:  
China Univ Geosci Wuhan, Natl Engn Res Ctr Geog Informat Syst, Wuhan 430078, Peoples R China  
Institution Address:  
China Univ Geosci Wuhan, Sch Comp Sci, Wuhan 430078, Peoples R China  
Institution Address:  
China State Shipbuilding Corp 722 Res Inst, Wuhan 430205, Peoples R China  
Institution Address:  
Wuhan Text Univ, Sch Comp Sci & Artificial Intelligence, Wuhan 430200, Peoples R China  
Institution Address:  
China Univ Geosci Wuhan, Hubei Key Lab Intelligent Geoinformat Proc, Wuhan 430078, Peoples R China  
Institution:  
Institution:  
Institution:  
China State Shipbuilding Corp 722 Res Inst  
Institution:  
Institution:

### 513. GRS-Det: An Anchor-Free Rotation Ship Detector Based on Gaussian-Mask in

摘要: Ship detection is a significant and challenging task in remote sensing. Due to the arbitrary-oriented property and large aspect ratio of ships, most of the existing detectors adopt rotation boxes to represent ships. However, manual-designed rotation anchors are needed in these detectors, which causes multiplied computational cost and inaccurate box regression. To address the abovementioned problems, an anchor-free rotation ship detector, named GRS-Det, is proposed, which mainly consists of a feature extraction network with selective concatenation module (SCM), a rotation Gaussian-Mask model, and a fully convolutional network-based detection module. First, a U-shape network with SCM is used to extract multiscale feature maps. With the help of SCM, the channel unbalance problem between different-level features in feature fusion is solved. Then, a rotation Gaussian-Mask is designed to model the ship based on its geometry characteristics, which aims at solving the mislabeling problem of rotation bounding boxes. Meanwhile, the Gaussian-Mask leverages context information to strengthen the perception of ships. Finally, multiscale feature maps are fed to the detection module for classification and regression of each pixel. Our proposed method, evaluated on ship detection benchmarks, including HRSC2016 and DOTA Ship data sets, achieves state-of-the-art results.  
Web of Science 核心合集中的 "被引频次":  
57  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
58  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000633493700055  
Indexed Date:  
2021-05-22  
Institution Address:  
Xidian Univ, Sch Artificial Intelligence, Xian 710071, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Sch Elect & Informat Engn, Xian 710049, Peoples R China  
Institution:  
Institution:

### 514. Accurate 3-D Semantic Segmentation of Point Clouds for Intelligent

摘要: Semantic segmentation based on LiDAR plays an important role in the environment perception, path planning, and decision control of unmanned ground systems. However, in challenging environments with complex background interleaving, the high proportion of drivable areas and significant target elements pose serious challenges to segmentation tasks. Existing studies attempt to address these challenges but still grapple with issues such as interclass semantic conflicts and inaccurate edge segmentation. To address these challenges effectively, a reliable point cloud segmentation network driven by multiview fusion and multiedge guidance is specially customized. Specifically, the network's point cloud multiview range image view (RIV) and bird's eye view (BEV) encoding backbones use lightweight pyramid architectures to extract specific details and semantic features of multiscale point clouds. Subsequently, a dedicated multiview edge guidance module focuses on enhancing interclass edge differentiation features from RIV and BEV, respectively. These features are fused with the encoding backbone hierarchy, and the extraction of edge information is supervised by a new hybrid edge loss to maximize the consistency of semantic segmentation and predicted edges. Furthermore, a multiview and multiscale feature fusion module based on a multihead attention mechanism is introduced to enhance the potential complementarity and interaction of the captured features. Through comprehensive experiments and ablation studies on the SemanticKITTI and RELLIS-3D datasets, recognized in the field of autonomous driving, the results demonstrate that our custom method is competitive in terms of intersection over union (IoU) accuracy and real-time indicators compared with typical baseline methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001295190900022  
Indexed Date:  
2024-10-10  
Institution Address:  
China Automot Engn Res Inst Co Ltd, Chongqing 401122, Peoples R China  
Institution Address:  
Jilin Univ, Sch Elect, Changchun 130012, Peoples R China  
Institution Address:  
Chinese Acad Agr Mechanizat Sci Grp Co Ltd, Beijing 100083, Peoples R China  
Institution Address:  
Shandong Univ Sci & Technol, Coll Transportat, Qingdao 266590, Peoples R China  
Institution:  
China Automot Engn Res Inst Co Ltd  
Institution:  
Institution:  
Chinese Acad Agr Mechanizat Sci Grp Co Ltd  
Institution:

### 515. AFENet: Attention Fusion Enhancement Network for Optic Disc Segmentation

摘要: Retinopathy of prematurity and ischemic brain injury resulting in periventricular white matter damage are the main causes of visual impairment in premature infants. Accurate optic disc (OD) segmentation has important prognostic significance for the auxiliary diagnosis of the above two diseases of premature infants. Because of the complexity and non-uniform illumination and low contrast between background and the target area of the fundus images, the segmentation of OD for infants is challenging and rarely reported in the literature. In this article, to tackle these problems, we propose a novel attention fusion enhancement network (AFENet) for the accurate segmentation of OD in the fundus images of premature infants by fusing adjacent high-level semantic information and multiscale low-level detailed information from different levels based on encoder-decoder network. Specifically, we first design a dual-scale semantic enhancement (DsSE) module between the encoder and the decoder inspired by self-attention mechanism, which can enhance the semantic contextual information for the decoder by reconstructing skip connection. Then, to reduce the semantic gaps between the high-level and low-level features, a multiscale feature fusion (MsFF) module is developed to fuse multiple features of different levels at the top of encoder by using attention mechanism. Finally, the proposed AFENet was evaluated on the fundus images of preterm infants for OD segmentation, which shows that the proposed two modules are both promising. Based on the baseline (Res34UNet), using DsSE or MsFF module alone can increase Dice similarity coefficients by 1.51 and 1.70%, respectively, whereas the integration of the two modules together can increase 2.11%. Compared with other state-of-the-art segmentation methods, the proposed AFENet achieves a high segmentation performance.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
1662-453X  
入藏号:  
WOS:000809820600001  
PubMed ID:  
35516802  
Indexed Date:  
2022-06-22

### 516. Continuous Sign Language Recognition With Multi-Scale Spatial-Temporal

摘要: Continuous Sign Language Recognition (CSLR) seeks to interpret the gestures used by people who are hard of hearing-mute individuals and translate them into natural language, thereby enhancing communication and interaction. A successful CSLR method relies on the continuous tracking of the presenter's gestures and facial movements. Existing CSLR methods struggle with fully leveraging fine-grained continuous frame information and often overlook the importance of multi-scale feature integration during decoding. To solve the above-mentioned issues, in this paper, we propose a spatial-temporal feature-enhanced network, called STNet for CSLR task. Firstly, for better continuous frame information exploration, based on the optimal transport algorithm, we first propose a spatial resonance module, which is used to extract the global common spatial features of two adjacent frames along the frame sequence. Secondly, we design a frame-wise loss to preserve and enhance the specific features of each frame. Lastly, to emphasize the multi-scale feature fusion, on the decoder side, we design a multi-temporal perception module, to allow each frame to focus on a larger range of other frames and enhance information interaction from different scales. Extensive experiments on three benchmark datasets including PHOENIX14, PHOENIX14-T, and CSL-Daily demonstrate that STNet consistently outperforms state-of-the-art methods, with a notable improvement of 2.9% in CSLR, showcasing its effectiveness and generalizability. Our approach provides a robust foundation for real-world applications such as sign language education and communication tools, while ablation and case studies highlight the impact of each module, paving the way for future research in CSLR.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001395266100010  
Indexed Date:  
2025-01-19  
Institution Address:  
Tokyo Inst Technol, Sch Informat & Commun Engn, Tokyo 1528550, Japan  
Institution Address:  
Univ Tokyo, Ctr Spatial Informat Sci, Tokyo 1138654, Japan  
Institution:  
Institution:

### 517. ScanGuard-YOLO: Enhancing X-ray Prohibited Item Detection with

摘要: To address the problem of low recall rate in the detection of prohibited items in X-ray images due to the severe object occlusion and complex background, an X-ray prohibited item detection network, ScanGuard-YOLO, based on the YOLOv5 architecture, is proposed to effectively improve the model's recall rate and the comprehensive metric F1 score. Firstly, the RFB-s module was added to the end part of the backbone, and dilated convolution was used to increase the receptive field of the backbone network to better capture global features. In the neck section, the efficient RepGFPN module was employed to fuse multiscale information from the backbone output. This aimed to capture details and contextual information at various scales, thereby enhancing the model's understanding and representation capability of the object. Secondly, a novel detection head was introduced to unify scale-awareness, spatial-awareness, and task-awareness altogether, which significantly improved the representation ability of the object detection heads. Finally, the bounding box regression loss function was defined as the WIOUv3 loss, effectively balancing the contribution of low-quality and high-quality samples to the loss. ScanGuard-YOLO was tested on OPIXray and HiXray datasets, showing significant improvements compared to the baseline model. The mean average precision (mAP@0.5) increased by 2.3% and 1.6%, the recall rate improved by 4.5% and 2%, and the F1 score increased by 2.3% and 1%, respectively. The experimental results demonstrate that ScanGuard-YOLO effectively enhances the detection capability of prohibited items in complex backgrounds and exhibits broad prospects for application.  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
eISSN:  
1424-8220  
入藏号:  
WOS:001140525200001  
PubMed ID:  
38202964  
Indexed Date:  
2024-02-02  
Institution Address:  
Yunnan Normal Univ, Sch Informat Sci & Technol, Kunming 650500, Peoples R China  
Institution:

### 518. Cross stage partial connections based weighted Bi-directional feature

摘要: Structural information is an essential component for efficient object detection. In many visual detection tasks, the objects with large structural deformation usually make up a large proportion. The shape, con-tour, and internal structure of the objects tend toward dramatic change, which easily causes troubles for efficient object detection. Therefore, how to detect these objects robustly and accurately is one of the sig-nificant challenges. To address this issue, we introduce a Cross Stage Partial connections-based weighted Bi-directional Feature Pyramid Network (CSP-BiFPN), which allows easy and efficient multi-scale feature fusion by cross-stage partial connections. Second, to enhance the model's spatial transformation capacity, the multi-scale feature maps extracted from the YOLO backbone network are processed by an enhanced spatial transformation network (ESTN) for spatial deformations. Based on these architectural modifica-tions and optimizations, we further develop a novel real-time robust object detection model called Bi-STN-YOLO. We evaluate the performance of the proposed method on four image datasets. The experi-mental results demonstrate that the proposed approach achieves significant improvements compared with the typical YOLO families and competitive performance compared to the state-of-the-arts in detec-tion tasks. (c) 2022 Elsevier B.V. All rights reserved.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
12  
被引频次合计:  
12  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000866415400007  
Indexed Date:  
2022-10-27  
Institution Address:  
Chinese Acad Sci, Inst Automation, State Key Lab Multimodal Artificial Intelligence S, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Artificial Intelligence, Beijing 100049, Peoples R China  
Institution Address:  
Nanchang Univ, Sch Informat Engn, Nanchang 330031, Peoples R China  
Institution Address:  
Huizhou Univ, Sch Elect Informat & Elect Engn, Huizhou 516007, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automation, State Key Lab Management & Control Complex Syst, Beijing 100190, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 519. Semantic Joint Monocular Remote Sensing Image Digital Surface Model

摘要: Digital surface model (DSM) presents height information of the Earth's surface and plays an important role in many remote sensing (RS) applications. Since the conventional acquisition of DSM is laborious and expensive, DSM reconstruction from monocular RS images has attracted extensive research in recent years, which is an ill-posed problem and thus rather challenging. Related works have achieved great accomplishments in this regard; however, they still face some limitations in training robustness, accuracy, and efficiency. To address the issues, a semantic joint monocular RS image DSM regression framework is proposed in this article, whose salient points include that: 1) semantic segmentation is integrated into the DSM regression task so that a shared backbone can extract complementary features from each objective to improve the performance of the individual task. Meanwhile, based on the consistency of the two training objectives, a two-stage joint loss function is introduced to improve the convergence and robustness of model training; 2) an encoding-decoding backbone is designed based on feature multiplexing, which simultaneously achieves multiscale feature fusion and information decoupling, thereby greatly reducing model parameters and improving efficiency while ensuring feature extraction effect; and 3) an iterative upsampling approach is introduced to transform the full-scale spatial features into large receptive-field and locally discriminative dynamic kernels, which are used to inpaint coarse-grained features while decoding, thus enhancing regression accuracy. Finally, experiments demonstrate the effectiveness of the proposal. It is easy to train and achieves superior or comparable accuracy compared with state-of-the-art related works while improving the efficiency by a large margin.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000805835300004  
Indexed Date:  
2022-06-17  
Institution Address:  
Beihang Hangzhou Innovat Inst Yuhang, Hangzhou 310023, Peoples R China  
Institution Address:  
Beihang Univ, Sch Automat Sci & Elect Engn, Beijing 100191, Peoples R China  
Institution:  
Beihang Hangzhou Innovat Inst Yuhang  
Institution:

### 520. Attention-based dual-path feature fusion network for automatic skin

摘要: Automatic segmentation of skin lesions is a critical step in Computer Aided Diagnosis (CAD) of melanoma. However, due to the blurring of the lesion boundary, uneven color distribution, and low image contrast, resulting in poor segmentation result. Aiming at the problem of difficult segmentation of skin lesions, this paper proposes an Attention-based Dual-path Feature Fusion Network (ADFFNet) for automatic skin lesion segmentation. Firstly, in the spatial path, a Boundary Refinement (BR) module is designed for the output of low-level features to filter out irrelevant background information and retain more boundary details of the lesion area. Secondly, in the context path, a Multi-scale Feature Selection (MFS) module is constructed for high-level feature output to capture multi-scale context information and use the attention mechanism to filter out redundant semantic information. Finally, we design a Dual-path Feature Fusion (DFF) module, which uses high-level global attention information to guide the step-by-step fusion of high-level semantic features and low-level detail features, which is beneficial to restore image detail information and further improve the pixel-level segmentation accuracy of skin lesion. In the experiment, the ISIC 2018 and PH2 datasets are employed to evaluate the effectiveness of the proposed method. It achieves a performance of 0.890/ 0.925 and 0.933 /0.954 on the F1-score and SE index, respectively. Comparative analysis with state-of-the-art segmentation methods reveals that the ADFFNet algorithm exhibits superior segmentation performance.  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1756-0381  
入藏号:  
WOS:001077369700001  
PubMed ID:  
37807076  
Indexed Date:  
2023-10-18  
Institution Address:  
Southwest Univ Sci & Technol, Sch Informat Engn, Mianyang, Peoples R China  
Institution Address:  
Southwest Univ Finance & Econ, Tianfu Coll, Mianyang, Peoples R China  
Institution Address:  
Southwest Univ Sci & Technol, Robot Technol Used Special Environm Key Lab Sichua, Mianyang 621010, Peoples R China  
Institution Address:  
Southwest Univ Sci & Technol, Sch Mfg Sci & Engn, Mianyang, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 521. MFEFNet: Multi-scale feature enhancement and Fusion Network for polyp

摘要: The polyp segmentation technology based on computer-aided can effectively avoid the deterioration of polyps and prevent colorectal cancer. To segment the polyp target precisely, the Multi-Scale Feature Enhancement and Fusion Network (MFEFNet) is proposed. First of all, to balance the network's predictive ability and complexity, ResNet50 is designed as the backbone network, and the Shift Channel Block (SCB) is used to unify the spatial location of feature mappings and emphasize local information. Secondly, to further improve the network's feature-extracting ability, the Feature Enhancement Block (FEB) is added, which decouples features, reinforces features by multiple perspectives and reconstructs features. Meanwhile, to weaken the semantic gap in the feature fusion process, we propose strong associated couplers, the Multi-Scale Feature Fusion Block (MSFFB) and the Reducing Difference Block (RDB), which are mainly composed of multiple cross-complementary information interaction modes and reinforce the long-distance dependence between features. Finally, to further refine local regions, the Polarized Self-Attention (PSA) and the Balancing Attention Module (BAM) are introduced for better exploration of detailed information between foreground and background boundaries. Experiments have been conducted under five benchmark datasets (Kvasir-SEG, CVC-ClinicDB, CVC-ClinicDB, CVC300 and CVC-ColonDB) and compared with state-of-the-art polyp segmentation algorithms. The experimental result shows that the proposed network improves Dice and mean intersection over union (mIoU) by an average score of 3.4% and 4%, respectively. Therefore, extensive experiments demonstrate that the proposed network performs favorably against more than a dozen state-of-the-art methods on five popular polyp segmentation benchmarks.  
Web of Science 核心合集中的 "被引频次":  
9  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
被引频次合计:  
9  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:000989941300001  
PubMed ID:  
36965326  
Indexed Date:  
2023-05-24  
Institution Address:  
Changchun Univ, Grad Sch, Changchun 130022, Jilin, Peoples R China  
Institution Address:  
Changchun Univ, Sch Elect Informat Engn, Changchun 130022, Jilin, Peoples R China  
Institution:  
Institution:

### 522. SFFEF-YOLO: Small object detection network based on fine-grained feature

摘要: Unmanned aerial vehicles (UAVs) images object detection has emerged as a research hotspot, yet remains significant challenge due to variable target scales and the high proportion of small objects caused by UAVs' diverse altitudes and angles. To address these issues, we propose a novel Small Object Detection Network Based on Fine-Grained Feature Extraction and Fusion(SFFEF-YOLO). First, we introduce a tiny prediction head to replace the large prediction head, enhancing the detection accuracy for tiny objects while reducing model complexity. Second, we design a Fine-Grained Information Extraction Module (FIEM) to replace standard convolutions. This module improves feature extraction and reduces information loss during downsampling by utilizing multi-branch operations and SPD-Conv. Third, we develop a Multi-Scale Feature Fusion Module (MFFM), which adds an additional skip connection branch based on the bidirectional feature pyramid network (BiFPN) to preserve fine-grained information and improve multi-scale feature fusion. We evaluated SFFEFYOLO on the VisDrone2019-DET and UAVDT datasets. Compared to YOLOv8, experimental results demonstrate that SFFEF-YOLO achieves a 9.9% mAP0.5 improvement on the VisDrone2019-DET dataset and a 3.6% mAP0.5 improvement on the UAVDT dataset.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0262-8856  
eISSN:  
1872-8138  
入藏号:  
WOS:001440806900001  
Indexed Date:  
2025-03-13  
Institution Address:  
Henan Univ, Sch Software, Kaifeng 475000, Peoples R China  
Institution Address:  
Henan Univ, Henan Prov Engn Res Ctr Intelligent Data Proc, Kaifeng 475000, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Engn Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:  
Institution:

### 523. Multidimensional electric power parameter time series forecasting and

摘要: With the advancement of big data and artificial intelligence, the intelligence level of power systems has significantly increased. Accurate prediction of power parameter time series fluctuations is crucial for the stable operation and efficient management of these systems. Traditional methods face challenges in capturing longterm dependencies and multi-scale features of multi-dimensional variables in nonlinear, non-stationary time series, limiting prediction accuracy. To address this, we propose an Adaptive Feature Fusion Convolution and Global-Local Dynamic Attention Reinforcement Learning (AFFC-GLDA-RL) model. The AFFC module enables multi-scale feature fusion and extraction from multi-dimensional variables, while the GLDA mechanism and RL strategy provide dynamic feature weighting and adaptive optimization, enhancing model focus on key features and robustness to noise. Additionally, a z-score-based anomaly detection method is incorporated to identify and analyze anomalous fluctuations in power parameter sequences. Experimental results demonstrate the model's superior performance in forecasting periodic power parameters, achieving R2 values above 0.99 and RMSE below 0.012. For more volatile parameters, the model maintains high accuracy with R2 above 0.85 and RMSE below 0.06. Furthermore, the model demonstrates strong predictive accuracy for 24-h ahead power parameter forecasts, providing robust support for stable operation and load forecasting in power systems.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0360-5442  
eISSN:  
1873-6785  
入藏号:  
WOS:001384519100001  
Indexed Date:  
2024-12-31  
Institution Address:  
Zhengzhou Univ Light Ind, Sch Bldg Environm Engn, Zhengzhou 450001, Peoples R China  
Institution Address:  
Zhengzhou Univ Light Ind, Zhengzhou Key Lab Elect Power Fire Safety, Zhengzhou 450001, Peoples R China  
Institution:  
Institution:

### 524. A lightweight multi-feature fusion network for unmanned aerial vehicle

摘要: UAV (Unmanned Aerial Vehicle) infrared object detection is crucial in pedestrian monitoring and traffic dispatch, which detects and locates objects in infrared images. In light of issues such as unnoticeable texture features and limited resolution of infrared image objects, a lightweight multi-scale feature fusion method for UAV infrared object detection is presented to enhance the performance of UAVs carrying intelligent devices to detect infrared objects. By changing the anchorless frame strategy of the YOLOX method, a lightweight MultiFeature Fusion Network (MFFNet) for UAV infrared ray (IR) image object detection is proposed. First, a lightweight backbone network is built using ShuffleNetv2\_block, spatial pyramid pooling, and other modules to reduce the network's number of parameters and inference time while maintaining its capacity to extract features. Second, we develop a multi-feature fusion module to improve the detection capabilities of the model for IR objects by fusing the local features and the overall characteristics of IR objects since the texture features of IR objects are challenging to employ, but the boundary information is evident. The boundary frame regression loss is then optimized using SCYLLA-IoU (SIoU) by comparing the predicted frame to the actual frame in terms of angle, distance, shape, and IoU (Intersection over Union), which forces the model to reach the optimum predicted box more quickly. The experimental results demonstrate that our method achieves an 81.5% mean average precision (mAP) with 4.21M parameters and an inference time of only 4.84ms per image, outperforming most networks in speed and accuracy.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1110-9823  
eISSN:  
2090-2476  
入藏号:  
WOS:001217798800001  
Indexed Date:  
2024-05-16  
Institution Address:  
Guizhou Univ, Coll Big Data & Informat Engn, Guiyang 550025, Guizhou, Peoples R China  
Institution Address:  
Guizhou Univ, State Key Lab Publ Big Data, Guiyang 550025, Guizhou, Peoples R China  
Institution Address:  
Guizhou Univ, Chongli Bldg, Guiyang 550025, Guizhou, Peoples R China  
Institution:  
Institution:  
Institution:

### 525. An Infrared Object Detection Method Based on Cross-domain Fusion Network

摘要: Because the infrared image lacks certain texture information,most target detection networks cannot achieve great detection results for infrared images. This paper proposes a cross-domain fusion network structure that combines multiple modal for infrared target detection. Using image conversion network without pairing,modal conversion of existing infrared dataset to generate a pseudo-visible light dataset. Then, this paper proposes a dual-channel multi-scale feature fusion structure in the infrared domain and the pseudo-visible light domain,uses feature pyramid network to obtain the feature map of each mode,and performs dual-modal feature fusion for multi-scale features. Finally,in order to make up for the lack of texture in the fusion process,this paper proposes a soft weight distribution module. By splicing the parameterized source domain,target domain and fusion domain features,the network weight is assigned and optimized through learning,thereby improving the accuracy of feature extraction and target detection. The experimental results show that the method in this paper has better infrared target detection performance compared with the conventional method.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
ISSN:  
1004-4213  
入藏号:  
WOS:000753173100031  
Indexed Date:  
2022-02-19  
Institution Address:  
Shanghai Maritime Univ, Coll Informat Engn, Shanghai 201306, Peoples R China  
Institution Address:  
Chinese Acad Sci, Shanghai Inst Tech Phys, Key Lab Intelligent Infrared Percept, Shanghai 200083, Peoples R China  
Institution:  
Institution:

### 526. A New Dual-Branch Network With Global Information for the Surface Defect

摘要: Photovoltaic (PV) wafers are a basic material for solar cells, and surface defect detection for solar PV wafers is necessary to improve the quality of solar PV. However, the complexity of the wafer manufacturing environment causes the surface defects of PV wafers to appear diverse, and it is still challenging in practical applications. In this research, a deep-learning (DL) network based on global information branches and dual-branch feature fusion (DBFF) is proposed for surface defect detection (SDD), named GFYolov7. First, for the issue of the low contrast on the wafer surface samples, the global information branching (GIB) is generated using the multiscale features of the backbone, which improves the information utilization of the backbone. Second, a DBFF module is designed for the GIB module, and a lightweight attention mechanism is used for the fused multiscale features, which serve to highlight defective regions and suppress irrelevant regions. The experimental results of GFYolov7 on NEU-DET and the GC10-DET reach 79.5 and 72.1 mAP and that on the real-world PV wafer dataset is 72.7 mAP. By comparing with other DL methods, GFYolov7 has been demonstrated that it has good detection ability and its detection speed can reach 61 FPS.  
Web of Science 核心合集中的 "被引频次":  
9  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
9  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001197673400186  
Indexed Date:  
2024-06-14  
Institution Address:  
China Univ Geosci, Sch Mech Engn & Elect Informat, Wuhan 430074, Peoples R China  
Institution Address:  
Guizhou Univ, Key Lab Adv Mfg Technol, Minist Educ, Guiyang 550025, Peoples R China  
Institution Address:  
China Univ Geosci, Shenzhen Res Inst, Shenzhen 518057, Peoples R China  
Institution Address:  
Huazhong Univ Sci & Technol, State Key Lab Digital Mfg Equipment & Technol, Wuhan 430074, Peoples R China  
Institution Address:  
Wuhan Reg Climate Ctr, Wuhan 430074, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Wuhan Reg Climate Ctr

### 527. Lightweight vehicle object detection network for unmanned aerial

摘要: Due to the limited computing power of unmanned aerial vehicles (UAVs) and the problems of missed detection and wrong detection of small objects, the current object detection algorithm cannot achieve real-time and high-precision detection. To solve these problems, we propose a vehicle detection network Shuffle CarNet for UAVs aerial images, which is composed of a feature extraction network, a feature fusion network, and a three-scale prediction network. First, according to the limited hardware resources of embedded devices, a lightweight feature extraction network Light CarNet is proposed by fusing the attention mechanism. Second, a four-scale feature bidirectional weighted fusion module is designed. According to the characteristics of the object scale, multilevel feature map bidirectional weighted fusion is selected for target classification and bounding box regression on three scales. Finally, Car-non-maximum suppression is used to reduce false detection and missed detection. Experiments show that compared with other algorithms on the VisDrone-2019 dataset, the proposed method improves the mean average precision by 1.14%, achieves a precision of 82.96%, and can meet the needs of real-time vehicle detection. The superiority of this method is proved by many comparative experiments.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:000965365700036  
Indexed Date:  
2023-04-30  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao, Peoples R China  
Institution Address:  
Tongji Univ, Dept Comp Sci & Technol, Shanghai, Peoples R China  
Institution:  
Institution:

### 528. SiamMAN: Siamese Multi-Phase Aware Network for Real-Time Unmanned Aerial

摘要: In this paper, we address aerial tracking tasks by designing multi-phase aware networks to obtain rich long-range dependencies. For aerial tracking tasks, the existing methods are prone to tracking drift in scenarios with high demand for multi-layer long-range feature dependencies such as viewpoint change caused by the characteristics of the UAV shooting perspective, low resolution, etc. In contrast to the previous works that only used multi-scale feature fusion to obtain contextual information, we designed a new architecture to adapt the characteristics of different levels of features in challenging scenarios to adaptively integrate regional features and the corresponding global dependencies information. Specifically, for the proposed tracker (SiamMAN), we first propose a two-stage aware neck (TAN), where first a cascaded splitting encoder (CSE) is used to obtain the distributed long-range relevance among the sub-branches by the splitting of feature channels, and then a multi-level contextual decoder (MCD) is used to achieve further global dependency fusion. Finally, we design the response map context encoder (RCE) utilizing long-range contextual information in backpropagation to accomplish pixel-level updating for the deeper features and better balance the semantic and spatial information. Several experiments on well-known tracking benchmarks illustrate that the proposed method outperforms SOTA trackers, which results from the effective utilization of the proposed multi-phase aware network for different levels of features.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
eISSN:  
2504-446X  
入藏号:  
WOS:001131042000001  
Indexed Date:  
2024-01-06  
Institution Address:  
Chinese Acad Sci, Changchun Inst Opt Fine Mech & Phys CIOMP, Changchun 130033, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Coll Mat Sci & Optoelect Technol, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 529. AgeDETR: Attention-Guided Efficient DETR for Space Target Detection

摘要: Recent advancements in space exploration technology have significantly increased the number of diverse satellites in orbit. This surge in space-related information has posed considerable challenges in developing space target surveillance and situational awareness systems. However, existing detection algorithms face obstacles such as complex space backgrounds, varying illumination conditions, and diverse target sizes. To address these challenges, we propose an innovative end-to-end Attention-Guided Encoder DETR (AgeDETR) model, since artificial intelligence technology has progressed swiftly in recent years. Specifically, AgeDETR integrates Efficient Multi-Scale Attention (EMA) Enhanced FasterNet block (EF-Block) within a ResNet18 (EF-ResNet18) backbone. This integration enhances feature extraction and computational efficiency, providing a robust foundation for accurately identifying space targets. Additionally, we introduce the Attention-Guided Feature Enhancement (AGFE) module, which leverages self-attention and channel attention mechanisms to effectively extract and reinforce salient target features. Furthermore, the Attention-Guided Feature Fusion (AGFF) module optimizes multi-scale feature integration and produces highly expressive feature representations, which significantly improves recognition accuracy. The proposed AgeDETR framework achieves outstanding performance metrics, i.e., 97.9% in mAP0.5 and 85.2% in mAP0.5:0.95, on the SPARK2022 dataset, outperforming existing detectors and demonstrating superior performance in space target detection.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:001324063100001  
Indexed Date:  
2024-10-07  
Institution Address:  
Chinese Acad Sci, Natl Space Sci Ctr, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution Address:  
Xidian Univ, Sch Telecommun Engn, Xian 710071, Peoples R China  
Institution:  
Institution:  
Institution:

### 530. MESR: Multistage Enhancement Network for Image Super-Resolution

摘要: Recently, the deep-learning-based image super-resolution methods have achieved astounding advancement. Whereas most of these methods utilize features from the low-resolution image space exclusively, and ignore the dependency between contextual features simultaneously, resulting in their limited ability to restore details. To this end, a multi-stage enhancement image network for super-resolution (MESR) is proposed. The network consists of two stages, where the first stage is used to generate a coarse reconstructed image, and the second one is to refine the coarse image, which enhances the super-resolution performance. Specifically, in the first stage, to acquire more abundant features, an effective funnel-like multi-scale feature extractor is proposed, incorporating a channel attention mechanism to boost the feature representation capability. Moreover, an adaptive weighted residual feature fusion block is designed to effectively explore and exploit the dependency between contextual features for generating more beneficial features. In the second stage, a refinement block is proposed to additionally strengthen the details of the reconstructed image by exploring the feature information from the high-resolution image space. Experimental results demonstrate that the proposed method achieves superior performance against the state-of-the-art SR methods in terms of both subjective visual quality and objective quantitative metrics.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:000800774000001  
Indexed Date:  
2022-06-06  
Institution Address:  
Huaqiao Univ, Coll Engn, Quanzhou 362021, Peoples R China  
Institution:

### 531. Borehole Depth Recognition Based on Improved YOLOX Detection

摘要: This study proposes a method for recognizing the drill depth in low-light underground environments, with the aim of addressing the issues of low efficiency and susceptibility to manual changes in the current methods. The method is based on an improved You Only Look Once X model. Initially, image data undergo enhancement and annotation. Secondly, it incorporates an attention mechanism to improve the feature extraction capability. The feature pyramid is utilized to minimize feature loss and facilitate better multi-scale feature fusion. Additionally, the loss function is optimized to enhance the localization ability of the prediction box. The enhanced model achieves an accuracy of 91.3$\%$, representing a 4.4$\%$ increase compared to the pre-improvement performance, and demonstrates improved positioning accuracy. Successful drilling depth measurements were carried out with the acquired positioning information.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0010-4620  
eISSN:  
1460-2067  
入藏号:  
WOS:001160410200001  
Indexed Date:  
2024-02-20  
Institution Address:  
Shandong Univ Sci & Technol, Coll Energy & Min Engn, Qingdao 266590, Peoples R China  
Institution:

### 532. A Multi-Stage Progressive Network with Feature Transmission and Fusion

摘要: Improving underwater image quality is crucial for marine detection applications. However, in the marine environment, captured images are often affected by various degradation factors due to the complexity of underwater conditions. In addition to common color distortions, marine snow noise in underwater images is also a significant issue. The backscatter of artificial light on marine snow generates specks in images, thereby affecting image quality, scene perception, and subsequently impacting downstream tasks such as target detection and segmentation. Addressing the issues caused by marine snow noise, we have designed a new network structure. In this work, a novel skip-connection structure called a dual channel multi-scale feature transmitter (DCMFT) is implemented to reduce information loss during downsampling in the feature encoding and decoding section. Additionally, in the feature transfer process for each stage, iterative attentional feature fusion (iAFF) modules are inserted to fully utilize marine snow features extracted at different stages. Finally, to further optimize the network's performance, we incorporate the multi-scale structural similarity index (MS-SSIM) into the loss function to ensure more effective convergence during training. Through experiments conducted on the Marine Snow Removal Benchmark (MSRB) dataset with an augmented sample size, our method has achieved significant results. The experimental results demonstrate that our approach excels in removing marine snow noise, with a peak signal-to-noise ratio reaching 38.9251 dB, significantly outperforming existing methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:001151132700001  
PubMed ID:  
38257449  
Indexed Date:  
2024-02-06  
Institution Address:  
Chinese Acad Sci, Inst Deep Sea Sci & Engn, Sanya 572000, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 533. A Fast and Accurate Human Pose Estimation Method Based on Multi-Scale

摘要: Human pose estimation (HPE) is a research hotspot in the field of computer vision. Most of the existing approaches first generate low-resolution representation from high-resolution representation through continuous serial downsampling, and then reconstruct high-resolution results from low-resolution features through continuous serial upsampling, which loses a lot of effective feature information and leads to slow model inference. In this paper, the Fast Accuracy Network (FANet), a framework that enables fast and high-accuracy HPE, is proposed. The innovation lies in that, first of all, a grid structure is proposed and adopted, which can be regarded as a set of deep paths and shallow paths. The structure uses multiple high-resolution and low-resolution branch pairs to perform skip-level connections at different scale-space levels so that the information can be exchanged between different resolution representations for many times. The feature information fusion of multi-scale space is realized to obtain more abundant feature information. Second, an improved bottleneck block is proposed to extract effective feature information with fewer parameters, ensuring that the computational burden is reduced without sacrificing accuracy performance. The experimental results show that, compared with other current models, FANet has faster inference speed on the premise of a slight improvement in accuracy performance.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0218-0014  
eISSN:  
1793-6381  
入藏号:  
WOS:000922652800001  
Indexed Date:  
2023-02-25  
Institution Address:  
Shanghai Maritime Univ, Dept Comp Sci & Technol, Shanghai 201306, Peoples R China  
Institution Address:  
Chinese Acad Sci, Shanghai Inst Microsyst & Informat Technol, State Key Lab Funct Mat Informat, Shanghai 200050, Peoples R China  
Institution Address:  
Chinese Acad Sci, Ctr Excellence Superconducting Elect, Shanghai 200050, Peoples R China  
Institution:  
Institution:  
Institution:

### 534. MSMA-Net: An Infrared Small Target Detection Network by Multiscale

摘要: Infrared small target detection plays a crucial role in various domains like early warning, national defense, and monitoring. Although existing detection methods have achieved some good results, they only rely on the original size and information of small targets for detection and are faced with challenges, such as the small size and obscure feature information of small targets. To overcome these limitations, this article introduces a coarse-to-fine detection network named MSMA-Net. This network initially determines the rough location of targets through a coarse preliminary screening, aiming to reduce false alarms and improve computational efficiency. Simultaneously, to improve the discriminability of the features and enhance the spatial details and resolution of the targets, the network utilizes multiscale super-resolution to transform low-resolution feature maps into high-resolution representations, gradually refining and strengthening the feature representation. Finally, the network employs a multilevel feature fusion attention mechanism to facilitate effective information transmission and fusion in multiscale and multilevel feature representations. This attention mechanism enhances the accuracy as well as robustness of object detection, ultimately obtaining accurate detection results. Extensive experimental results demonstrate that compared with existing detection methods, our approach can effectively suppress false alarms and get better performance even when the target has a small size and obscure feature information.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
7  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001136708200013  
Indexed Date:  
2024-02-04  
Institution Address:  
Zhengzhou Univ, Sch Elect & Informat Engn, Zhengzhou 450001, Peoples R China  
Institution:

### 535. Enhanced Multi-Scale Features Mutual Mapping Fusion Based on Reverse

摘要: Unsupervised anomaly detection methods based on knowledge distillation have exhibited promising results. However, there is still room for improvement in the differential characterization of anomalous samples. In this article, a novel anomaly detection and localization model based on reverse knowledge distillation is proposed, where an enhanced multi-scale feature mutual mapping feature fusion module is proposed to greatly extract discrepant features at different scales. This module helps enhance the difference in anomaly region representation in the teacher-student structure by inhomogeneously fusing features at different levels. Then, the coordinate attention mechanism is introduced in the reverse distillation structure to pay special attention to dominant issues, facilitating nice direction guidance and position encoding. Furthermore, an innovative single-category embedding memory bank, inspired by human memory mechanisms, is developed to normalize single-category embedding to encourage high-quality model reconstruction. Finally, in several categories of the well-known MVTec dataset, our model achieves better results than state-of-the-art models in terms of AUROC and PRO, with an overall average of 98.1%, 98.3%, and 95.0% for detection AUROC scores, localization AUROC scores, and localization PRO scores, respectively, across 15 categories. Extensive experiments are conducted on the ablation study to validate the contribution of each component of the model.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2332-7790  
入藏号:  
WOS:001270298100005  
Indexed Date:  
2024-07-24  
Institution Address:  
Univ Shanghai Sci & Technol, Sch Opt Elect Comp Engn, Shanghai 200093, Peoples R China  
Institution:

### 536. Single Image Super-Resolution Using Asynchronous Multi-Scale Network

摘要: An existing multi-scale residual network (MSRN) has demonstrated its success on conducting the single image super-resolution (SISR) task. The MSRN consists of a number of multi-scale residual blocks (MSRBs), and each MSRB performs convolutions by exploiting two different sizes of windows for conducting multi-scale feature extraction. The smaller window is used to extract image features at a low scale, while the larger one is used for a high scale. To significantly reduce the number of parameters involved in the MSRB, a new feature extraction module, called the asynchronous multi-scale block (AMB), is proposed in this paper. It is based on the fact that the larger window used in the MSRB can be replaced by two smaller windows without affecting the original MSRB's function. Consequently, by replacing each MSRB with our AMB, an asynchronous multi-scale network (AMNet) is then constructed, which can yield a significant reduction on computational complexity. This means that more AMBs can be used in our AMNet to deliver superior SISR performance, while maintaining the same or comparable computational complexity to that of the MSRN. To consolidate all image features generated from all scales, a new fusion scheme, called the adaptive feature fusion block (AFFB), is proposed that weights the extracted features according to their importance for further increasing SISR's performance. Extensive experimental results have clearly shown the superiority of our proposed AMNet when compared with multiple state-of-the-arts.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
被引频次合计:  
8  
ISSN:  
1070-9908  
eISSN:  
1558-2361  
入藏号:  
WOS:000697816900002  
Indexed Date:  
2021-09-29  
Institution Address:  
Soochow Univ, Sch Comp Sci & Technol, Suzhou 215006, Peoples R China  
Institution Address:  
Nanyang Technol Univ, Sch Elect & Elect Engn, Singapore 639798, Singapore  
Institution:  
Institution:

### 537. Point Cloud Feature Extraction Network Based on Multiscale Feature

摘要: Accurate feature extraction in point cloud registration is often hindered by noise, surface complexity, overlap, and scale differences, which limit improvements in registration. To address this issue, this study proposes a point cloud registration algorithm based on the dynamic fusion of multiscale features. First, by employing sparse convolution operations at different depths, multilevel scale feature information is extracted from the point cloud data, obtaining rich levels of detail from local and global structures. Subsequently, the multilevel scale features are concatenated to form a fused feature representation, which enhances the integrity and accuracy of features. Additionally, the algorithm introduces a squeezeexcitation attention mechanism for the network skip connections to adaptively learn and reinforce important feature information. Concurrently, a global context module is integrated at the residual position to better capture global structural information. Finally, registration is completed by estimating the rigid transformation matrix through the random sample consensus (RANSAC) algorithm. Experimental results demonstrate significant advantages in feature extraction and registration accuracy compared to mainstream methods, effectively improving the performance of point cloud registration.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001438403100010  
Indexed Date:  
2025-03-13  
Institution Address:  
North Univ China, Sch Data Sci & Technol, Taiyuan 030051, Shanxi, Peoples R China  
Institution Address:  
Shanxi Prov Key Lab Machine Vis & Virtual Real, Taiyuan 030051, Shanxi, Peoples R China  
Institution Address:  
North Automat Control Technol Inst, Dept Simulat Equipment, Taiyuan 030006, Shanxi, Peoples R China  
Institution:  
Institution:  
Shanxi Prov Key Lab Machine Vis & Virtual Real  
Institution:  
North Automat Control Technol Inst

### 538. Surface defect detection method for discarded mechanical parts under

摘要: With a significant number of mechanical products approaching the retirement phase, the batch recycling of discarded mechanical parts necessitates a preliminary assessment of their surface condition. However, the presence of surface rust poses a challenge to defect identification. Therefore, this paper proposes a method for detecting heavily rusted surface defects based on an improved YOLOv8n network. In the Backbone, the C2f-DBB module of re-parameterized deep feature extraction was introduced, and the attention module was designed to improve the accuracy of information extraction. In the Neck part, a Bi-Afpn multiscale feature fusion strategy is designed to facilitate information exchange between features at different scales. Finally, Focal-CIoU is employed as the bounding box loss function to enhance the network's localization performance and accuracy for defects. Experimentally, it is proved that the improved network in this paper improves the Recall, Precision, and mAP0.5 by 1.2%, 2.1%, and 1.9%, respectively, on the original basis, which is better than other network models.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2045-2322  
入藏号:  
WOS:001197575500018  
PubMed ID:  
38575736  
Indexed Date:  
2024-04-30  
Institution Address:  
Wuhan Univ Sci & Technol, Key Lab Met Equipment & Control Technol, Minist Educ, Wuhan 430081, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Hubei Key Lab Mech Transmiss & Mfg Engn, Wuhan 430081, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Precis Mfg Inst, Wuhan 430081, Peoples R China  
Institution:  
Institution:  
Institution:

### 539. A Lightweight and Accurate Method for Detecting Traffic Flow in Real

摘要: Traffic flow detection provides significant information for intelligent transportation systems. However, as the mainstream research direction, vision-based traffic flow detection methods currently face the challenges of a trade-off between accuracy and speed. Furthermore, it is crucial that modularization be incorporated into the system design process to enhance the maintainability and flexibility of the system. To achieve this, we propose a modular design method that divides this task into three parts: vehicle detecting, vehicle tracking, and vehicle counting. As an important link of the system, vehicle detection greatly influences the accuracy and speed of the system. We therefore introduce a lightweight network called feature adaptive fusionYOLOX, which is based on YOLOX. Specifically, in order to eliminate redundant information brought by bilinear interpolation, we propose a feature-level upsampling method called channel to spatial, which enables upsampling without additional calculations. Based on this module, we design a lightweight, multiscale feature fusion module, feature adaptive fusion pyramid network (FAFPN). Compared with PA-FPN, tal tests, the traffic flow detection method proposed in detection in real time.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1343-0130  
eISSN:  
1883-8014  
入藏号:  
WOS:001110853100003  
Indexed Date:  
2024-01-07  
Institution Address:  
Beijing Inst Technol, Sch Automat, 5 South St, Beijing 100081, Peoples R China  
Institution:

### 540. Lightweight Multi-Scale Feature Fusion Network for Salient Object

摘要: Salient object detection in optical remote sensing images (ORSI-SOD) encounters notable challenges, mainly because of the small scale of salient objects and the similarity between these objects and their backgrounds in images captured by satellite and aerial sensors. Conventional approaches frequently struggle to efficiently leverage multi-scale and multi-stage features. Moreover, these methods usually rely on sophisticated and resource-heavy architectures, which can limit their practicality and efficiency in real-world applications. To overcome these limitations, this paper proposes a novel lightweight network called the Multi-scale Feature Fusion Network (MFFNet). Specifically, a Multi-stage Information Fusion (MIF) module is created to improve the detection of salient objects by effectively integrating features from multiple stages and scales. Additionally, we design a Semantic Guidance Fusion (SGF) module to specifically alleviate the problem of semantic dilution often observed in U-Net architecture. Comprehensive evaluations on two benchmark datasets show that the MFFNet attains outstanding performance in four out of eight evaluation metrics while only having 12.14M parameters and 2.75G FLOPs. These results highlight significant advancements over 31 state-of-the-art models, underscoring the efficiency of MFFNet in salient object-detection tasks.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001393572600001  
Indexed Date:  
2025-01-15  
Institution Address:  
Chengdu Univ Technol, Coll Comp Sci & Cyber Secur, Chengdu 610059, Peoples R China  
Institution:

### 541. Wheel Tread Anomaly Detection Based on Attentional Reverse Knowledge

摘要: Wheels are an essential part of railway trains; thus, defects on the wheel tread present serious risk regarding the safety of railway trains. Due to the limited samples of wheel tread defects in practice, the corresponding supervised detection model is insufficient. To solve this problem, an unsupervised knowledge distillation anomaly detection model is proposed to detect wheel tread anomalies. Accordingly, UNet is employed to segment the tread region and reduce the influence of non-tread regions on the anomaly detection model. An attention mechanism is then added after the multiscale feature fusion to improve the ability of the student network to reconstruct normal features in the reverse knowledge distillation structure, as well as enhance the reconstruction of normal features. From the experimental results, the improved model achieves the performance indexes of 93.8% area under receiver operating characteristic curve, 82.3% precision, 95.4% recall, and 87.0% accuracy considering the railway wheel tread dataset. Compared with the original model, the detection performance of the model is improved.  
摘要:  
车轮是铁路列车走行部的重要部件,车轮踏面上产生的缺陷严重危害着铁路列车的安全运行。由于实际中车轮踏面缺陷样本有限,有监督检测模型对缺陷的检测不具有鲁棒性。针对此问题,提出使用无监督的知识蒸馏异常检测模型实现对车轮踏面的异常检测任务。首先,使用UNet对踏面区域进行分割,减少非踏面区域对异常检测模型的影响;然后,在多尺度特征聚合之后添加一个注意力机制,提升反向知识蒸馏结构中学生网络对正常特征的重建能力,增强学生网络对正常特征重建的效果。实验结果表明:在铁路车轮踏面数据集上,改进后的模型能够达到93.8%的受试者工作特性曲线下的面积、82.3%的精准率、95.4%的召回率、87.0%的准确率。与原模型相比,改进后的模型检测性能得到提升。  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7642630  
Indexed Date:  
2024-03-29  
Institution Address:  
西南交通大学物理科学与技术学院, 成都, 四川 610000, 中国  
Institution Address:  
School of Physical Science and Technology, Southwest Jiaotong University, Chengdu, Sichuan 610000, China  
Institution:  
西南交通大学物理科学与技术学院  
Institution:  
School of Physical Science and Technology, Southwest Jiaotong University

### 542. Enhancing YOLOv8 for Improved Instance Segmentation of Automotive

摘要: To address the shortcomings of manual damage assessment and issues with conventional vehicle damage detection models in the context of intelligent vehicles,it proposes EIS-YOLO,an enhanced instance segmentation model based on YOLOv8.It introduces CRDB,a novel multi-scale feature fusion and channel reduction module that replaces C2f,reducing parameters by 20.15% while improving fusion efficiency.Additionally,HRFPN structure maintains high-resolution branches,facilitates finer detail and semantic exchange,and includes AFF and BiAM attention modules for deeper feature integration.An efficient E-FPN and an extra output head are utilized to better identify small damages and edges.Evaluated on CarDD dataset,CRDB improves multi-task accuracy by 2 percentage points,and the integrated EIS-YOLO model with HRFPN sees a 4.4 percentage points boost in [PB] and 6.6 percentage points in [PM] over the baseline,all while maintaining a lighter weight and lower computational complexity.  
摘要:  
针对人工定损方式无法满足智能汽车时代的发展要求,及传统汽车伤损检测模型精度低、信息少、难部署等问题,提出了改进YOLOv8的汽车伤损实例分割模型EIS-YOLO。在主干网络中设计了一个多尺度特征融合与通道数减小的CRDB模块,取代传统C2f模块,显著减少了参数量的同时提高了特征融合的能力;提出了保留高分辨率分支的HRFPN结构,以加强细节信息保留能力,增强细节与语义信息的交换,该结构通过AFF和BiAM注意力融合模块增强了深层传递,经由简化冗余连接的E-FPN完成特征融合。还增加了一个额外的输出头捕捉细小伤损,提高了模型对小目标伤损及伤损边缘的精确识别。在CarDD数据集上,主干网络部分提出的CRDB模块对比C2f模块实现了同架构下计算量减小20.15%,同时多任务平均准确率提升2个百分点,在此基础上,结合HRFPN结构与额外输出头设计的模型整体的准确率[PB]、[PM]相较于基准模型分别提升了4.4和6.6个百分点,且模型更轻量,计算复杂度更低。  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1002-8331  
入藏号:  
CSCD:7794595  
Indexed Date:  
2024-11-01  
Institution Address:  
辽宁科技大学计算机与软件工程学院, 鞍山, 辽宁 114051, 中国  
Institution Address:  
School of Computer Science and Software Engineering,University of Science and Technology Liaoning, Anshan, Liaoning 114051, China  
Institution:  
辽宁科技大学计算机与软件工程学院  
Institution:  
School of Computer Science and Software Engineering,University of Science and Technology Liaoning

### 543. Multisource High-Resolution Remote Sensing Image Vegetation Extraction

摘要: High-resolution remote sensing image-based vegetation monitoring is a hot topic in remote sensing technology and applications. However, when facing large-scale monitoring across different sensors in broad areas, the current methods suffer from fragmentation and weak generalization capabilities. To address this issue, this paper proposes a multisource high-resolution remote sensing image-based vegetation extraction method that considers the comprehensive perception of multiple features. First, this method utilizes a random forest model to perform feature selection for the vegetation index, selecting an index that enhances the otherness between vegetation and other land features. Based on this, a multifeature synthesis perception convolutional network (MSCIN) is constructed, which enhances the extraction of multiscale feature information, global information interaction, and feature cross-fusion. The MSCIN network simultaneously constructs dual-branch parallel networks for spectral features and vegetation index features, strengthening multiscale feature extraction while reducing the loss of detailed features by simplifying the dense connection module. Furthermore, to facilitate global information interaction between the original spectral information and vegetation index features, a dual-path multihead cross-attention fusion module is designed. This module enhances the differentiation of vegetation from other land features and improves the network's generalization performance, enabling vegetation extraction from multisource high-resolution remote sensing data. To validate the effectiveness of this method, we randomly selected six test areas within Anhui Province and compared the results with three different data sources and other typical methods (NDVI, RFC, OCBDL, and HRNet). The results demonstrate that the MSCIN method proposed in this paper, under the premise of using only GF2 satellite images as samples, exhibits robust accuracy in extraction results across different sensors. It overcomes the rapid degradation of accuracy observed in other methods with various sensors and addresses issues such as internal fragmentation, false positives, and false negatives caused by sample generalization and image diversity.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:001168469200001  
Indexed Date:  
2024-03-02  
Institution Address:  
Anhui Univ, Sch Resources & Environm Engn, Hefei 230601, Peoples R China  
Institution Address:  
Anhui Univ, Stony Brook Inst, Hefei 230601, Peoples R China  
Institution Address:  
Anhui Univ, Inst Phys Sci & Informat Technol, Hefei 230601, Peoples R China  
Institution Address:  
Anhui Univ, Sch Artificial Intelligence, Hefei 230601, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 544. A Multi-Scale Feature Fusion Deep Learning Network for the Extraction of

摘要: In the face of global population growth and climate change, the protection and rational utilization of cropland are crucial for food security and ecological balance. However, the complex topography and unique ecological environment of the Qinghai-Tibet Plateau results in a lack of high-precision cropland monitoring data. Therefore, this paper constructs a high-quality cropland dataset for the YarlungZangbo-Lhasa-Nyangqv River region of the Qinghai-Tibet Plateau and proposes an MSC-ResUNet model for cropland extraction based on Landsat data. The dataset is annotated at the pixel level, comprising 61 Landsat 8 images in 2023. The MSC-ResUNet model innovatively combines multiscale features through residual connections and multiscale skip connections, effectively capturing features ranging from low-level spatial details to high-level semantic information and further enhances performance by incorporating depthwise separable convolutions as part of the feature fusion process. Experimental results indicate that MSC-ResUNet achieves superior accuracy compared to other models, with F1 scores of 0.826 and 0.856, and MCC values of 0.816 and 0.847, in regional robustness and temporal transferability tests, respectively. Performance analysis across different months and band combinations demonstrates that the model maintains high recognition accuracy during both growing and non-growing seasons, despite the study area's complex landforms and diverse crops.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001351988800001  
Indexed Date:  
2024-11-17  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 545. SOD-YOLO: Small-Object-Detection Algorithm Based on Improved YOLOv8 for

摘要: The rapid development of unmanned aerial vehicle (UAV) technology has contributed to the increasing sophistication of UAV-based object-detection systems, which are now extensively utilized in civilian and military sectors. However, object detection from UAV images has numerous challenges, including significant variations in the object size, changing spatial configurations, and cluttered backgrounds with multiple interfering elements. To address these challenges, we propose SOD-YOLO, an innovative model based on the YOLOv8 model, to detect small objects in UAV images. The model integrates the receptive field convolutional block attention module (RFCBAM) in the backbone network to perform downsampling, improving feature extraction efficiency and mitigating the spatial information sparsity caused by downsampling. Additionally, we developed a novel neck architecture called the balanced spatial and semantic information fusion pyramid network (BSSI-FPN) designed for multi-scale feature fusion. The BSSI-FPN effectively balances spatial and semantic information across feature maps using three primary strategies: fully utilizing large-scale features, increasing the frequency of multi-scale feature fusion, and implementing dynamic upsampling. The experimental results on the VisDrone2019 dataset demonstrate that SOD-YOLO-s improves the mAP50 indicator by 3% compared to YOLOv8s while reducing the number of parameters and computational complexity by 84.2% and 30%, respectively. Compared to YOLOv8l, SOD-YOLO-l improves the mAP50 indicator by 7.7% and reduces the number of parameters by 59.6%. Compared to other existing methods, SODA-YOLO-l achieves the highest detection accuracy, demonstrating the superiority of the proposed method.  
Web of Science 核心合集中的 "被引频次":  
10  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
10  
eISSN:  
2072-4292  
入藏号:  
WOS:001305440800001  
Indexed Date:  
2024-09-10  
Institution Address:  
Qilu Aerosp Informat Res Inst, Jinan 250132, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100190, Peoples R China  
Institution:  
Qilu Aerosp Informat Res Inst  
Institution:

### 546. CUFuse: Camera and Ultrasound Data Fusion for Rail Defect Detection

摘要: This paper proposes a multi-source data fusion algorithm for rail surface defect detection in both camera-based rail inspection images and ultrasound B-scan images. First, we design a rail surface segmentation algorithm based on image bilateral filtering, Sobel edge detection, and rail surface edge detection to extract the rail surface area. Second, we build a camera and ultrasound data fusion (CUFuse) model for rail surface defect detection, including two main networks: multi-source data feature extraction and multi-scale feature fusion networks. The multi-source data feature extraction network consists of two BoTNet 50 networks as feature extraction networks to extract five stages of features in camera-based images and ultrasound B-scan images. The multi-scale feature fusion network consists of five feature fusion modules to fuse the feature information output by the multi-source data feature extraction network. Finally, we use the CUFuse model to detect the rail surface defect dataset, and output five rail surface state types, including Light, Moderate, Severe, Normal, and Joint. The results show that the accuracy of the CUFuse model is 96.97%, which can accomplish the task of rail surface defect detection on railway sites.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
35  
被引频次合计:  
35  
ISSN:  
1524-9050  
eISSN:  
1558-0016  
入藏号:  
WOS:000829083500001  
Indexed Date:  
2022-07-31  
Institution Address:  
Southwest Jiaotong Univ, MOE Key Lab High Speed Railway Engn, Minist Educ, Chengdu 610031, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, Sch Civil Engn, Chengdu 610031, Peoples R China  
Institution Address:  
Shanghai Dongfang Maritime Engn Technol Co Ltd, Shanghai 200011, Peoples R China  
Institution Address:  
China Railway First Survey & Design Inst Grp Co L, Xian 710043, Peoples R China  
Institution:  
Institution:  
Institution:  
Shanghai Dongfang Maritime Engn Technol Co Ltd  
Institution:  
China Railway First Survey & Design Inst Grp Co L

### 547. Multiscale Modality-Similar Learning Guided Weakly Supervised RGB-T

摘要: With the development of sensor technology and its numerous applications in intelligent surveillance systems, RGB-thermal (RGB-T) cross-modal crowd counting uses data from different sensors as source data and has received extensive attention from academia and industry. From the feature extraction aspect, the existing cross-modal methods mainly adopt multiple parallel large convolution kernels for the notable crowd-scale variation problem, resulting in a large number of parameters. From the supervision aspect, the existing cross-modal crowd-counting methods adopt a fully supervised framework, and it requires time-consuming and laborious pixel-level supervision. In this regard, this article proposes a multiscale modality-similar guided weakly supervised cross-modal crowd-counting method, including a designed multiscale context-level feature fusion (MCFF) module and a modality-similar weakly supervised framework. In particular, the proposed multiscale module decouples the square convolution in different directions equivalently to solve the problems of feature redundancy and parameter increase. The proposed weakly supervised framework explores the similarity of cross-modal crowd semantic features to bootstrap the model with only image-level supervised information. Experimental results on two public RGB-T benchmarks, one RGB-D benchmark, and the collected real-world data show that the proposed weakly supervised method can achieve counting accuracy competitive with existing representative fully supervised methods. The extensive ablation studies validate the positive gain of the core modules on the final counting performance improvement.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001379128100067  
Indexed Date:  
2024-12-28  
Institution Address:  
Yanshan Univ, Sch Informat Sci & Engn, Qinhuangdao 066004, Hebei, Peoples R China  
Institution Address:  
Yanshan Univ, Key Lab Comp Virtual Technol & Syst Integrat Hebe, Qinhuangdao 066004, Hebei, Peoples R China  
Institution Address:  
Yanshan Univ, Hebei Key Lab Software Engn, Qinhuangdao 066004, Hebei, Peoples R China  
Institution:  
Institution:  
Institution:

### 548. ISLS: An Illumination-Aware Sauce-Packet Leakage Segmentation Method

摘要: The segmentation of abnormal regions is vital in smart manufacturing. The blurring sauce-packet leakage segmentation task (BSLST) is designed to distinguish the sauce packet and the leakage's foreground and background at the pixel level. However, the existing segmentation system for detecting sauce-packet leakage on intelligent sensors encounters an issue of imaging blurring caused by uneven illumination. This issue adversely affects segmentation performance, thereby hindering the measurements of leakage area and impeding the automated sauce-packet production. To alleviate this issue, we propose the two-stage illumination-aware sauce-packet leakage segmentation (ISLS) method for intelligent sensors. The ISLS comprises two main stages: illumination-aware region enhancement and leakage region segmentation. In the first stage, YOLO-Fastestv2 is employed to capture the Region of Interest (ROI), which reduces redundancy computations. Additionally, we propose image enhancement to relieve the impact of uneven illumination, enhancing the texture details of the ROI. In the second stage, we propose a novel feature extraction network. Specifically, we propose the multi-scale feature fusion module (MFFM) and the Sequential Self-Attention Mechanism (SSAM) to capture discriminative representations of leakage. The multi-level features are fused by the MFFM with a small number of parameters, which capture leakage semantics at different scales. The SSAM realizes the enhancement of valid features and the suppression of invalid features by the adaptive weighting of spatial and channel dimensions. Furthermore, we generate a self-built dataset of sauce packets, including 606 images with various leakage areas. Comprehensive experiments demonstrate that our ISLS method shows better results than several state-of-the-art methods, with additional performance analyses deployed on intelligent sensors to affirm the effectiveness of our proposed method.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:001231642500001  
PubMed ID:  
38794069  
Indexed Date:  
2024-05-31  
Institution Address:  
Nanjing Univ Posts & Telecommun, Sch Internet Things, Nanjing 210023, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Sch Comp Sci, Nanjing 210023, Peoples R China  
Institution Address:  
Natl Univ Def Technol, Res Inst 63, Nanjing 210007, Peoples R China  
Institution Address:  
Northern Informat Control Res Acad Grp Co Ltd, Nanjing 211153, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Northern Informat Control Res Acad Grp Co Ltd

### 549. Multi-Scale Feature Enhancement for Saliency Object Detection Algorithm

摘要: Aimed at existing saliency object detection models with problems of front and back view misclassification and edge blur, this study proposes an algorithm with multi-scale feature enhancement. In this algorithm, the feature maps of salient objects are extracted using VGG16. Multi-scale Feature Fusion Module is added to enhance the detailed information of the second feature layer and the semantic information of the fifth feature layer, which effectively improves the characterization ability of the second feature layer on the edges of salient objects and the fifth feature layer on salient objects. Simultaneously, Feature Enhancement Fusion Module is added to achieve the full fusion of local detail information and global semantic information through layer-by-layer fusion from deep to shallow, which is used to obtain a feature map with complete feature information. Finally, a complete prediction map with clear edges is obtained by training the network model. The performance of the proposed algorithm is compared with six algorithms, Amulet, R3Net, PoolNet, MINet, PurNet, and NSAL, on the HKU-IS, ECSSD, DUT-OMRON, and DUTS-TE datasets. MAE (Mean Absolute Error) values were decreased by 0.011, 0.009, 0, -0.001, 0.001, 0.003. F-measure were improved by 0.037, 0.019, 0.013, 0.017, 0.015, 0.09. E-measure were improved by: null, -0.008, 0.003, 0.005, -0.014, 0.047. S-measure were improved by: 0.073, 0.041, 0.016, 0.021, 0.016, 0.101. Compared with existing algorithms, the proposed algorithm can obtain better detection results and accurately identify all regions of significant objects.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001080860000001  
Indexed Date:  
2023-10-26  
Institution Address:  
Yancheng Inst Technol, Sch Informat Technol, Yancheng 224051, Peoples R China  
Institution Address:  
Yancheng XiongYing Precis Machinery Co Ltd, Yancheng 224006, Peoples R China  
Institution:  
Institution:  
Yancheng XiongYing Precis Machinery Co Ltd

### 550. Improved PointNet with accuracy and efficiency trade-off for online

摘要: Structural light vision sensing is widely applied in online detecting defects of laser processing due to their antiinterference ability for laser beams. However, the existing algorithms cannot extract the characteristics of weld defects with high precision. The computing cost of large-scale point cloud data is high. The balance between them is the main challenge to achieve online detection. To improve accuracy and reduce computation costs, this study uses point cloud data with depth information and proposes a point cloud segmentation method. It is a novelty method based on PointNet framework that has been verified for laser welding defect detection. Specifically, it used the PointNet framework as the backbone. It extracted enough local features of weld defects by multi-scale feature fusion, which concatenated features from different feature extraction layers to learn enough features to improve detection accuracy. The experiments were conducted on the real dataset of welds. The results showed its competitive performance in weld bead measurement and classification segmentation, and the accuracy of this method is 97.4 %. The proposed method improved mean intersection-over-union (mIoU) by 2.1 % compared with its backbone (PointNet), indicating a better segmentation accuracy. In addition, the proposed method improved detection speed compared with PointNet++. It can reach 60 frames per second, 7.5 times faster than PointNet++ and meet the online monitoring requirements. To conclude, the new detection method based on the improved PointNet with higher accuracy and faster speed of detection has a wide application prospect thanks to its novel model.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0143-8166  
eISSN:  
1873-0302  
入藏号:  
WOS:001327945500001  
Indexed Date:  
2024-10-12  
Institution Address:  
Wuhan Univ, Inst Technol Sci, Wuhan 430072, Peoples R China  
Institution Address:  
Wuhan Univ, Dong Fureng Inst Econ & Social Dev, Wuhan 430072, Peoples R China  
Institution Address:  
Wuhan Qingchuan Univ, Sch Mech & Elect Engn, Wuhan 430204, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Power & Mech Engn, Wuhan 430072, Peoples R China  
Institution:  
Institution:  
Institution:  
Wuhan Qingchuan Univ  
Institution:

### 551. Object Detection in Optical Remote Sensing Images Based on FFC-SSD Model

摘要: For the applications of efficient high-precision object detection in optical remote sensing (RS) images, this paper focuses on the difficulty of improving the detection accuracy of the SSD (single shot multibox detector) model on small and densely distributed objects in such images. An improved model FFC-SSD (multi-scale feature fusion & clustering SSD) is thereby proposed. For this purpose, a bounding-box group clustering (BGC) module is designed. Group clustering is implemented to obtain default object frame parameters that are more consistent with the size distribution of object samples and gives more attention to small objects. This module effectively improves the networks ability to extract object locations. Then, an efficient de-pooling multi-scale feature fusion (MSFF) module is designed to enhance the ability of the model to extract object features and effectively reduce the efficiency loss of the model at the same time. The experimental results demonstrate the effectiveness and applicability of the FFC-SSD model for object detection in optical remote sensing images. The proposed model achieves a favorable balance between precision and efficiency and has high detection accuracy on small objects.  
Web of Science 核心合集中的 "被引频次":  
10  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
5  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
15  
ISSN:  
0253-2239  
入藏号:  
WOS:000822427600013  
Indexed Date:  
2022-07-21  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Key Lab Quantitat Remote Sensing Informat Technol, Beijing 100094, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 552. Multi-Scale Feature Fusion Network for Low-Dose CT Denoising

摘要: Computed tomography (CT) is an imaging technique extensively used in medical treatment, but too much radiation dose in a CT scan will cause harm to the human body. Decreasing the dose of radiation will result in increased noise and artifacts in the reconstructed image, blurring the internal tissue and edge details. To get high-quality CT images, we present a multi-scale feature fusion network (MSFLNet) for low-dose CT (LDCT) denoising. In our MSFLNet, we combined multiple feature extraction modules, effective noise reduction modules, and fusion modules constructed using the attention mechanism to construct a horizontally connected multi-scale structure as the overall architecture of the network, which is used to construct different levels of feature maps at all scales. We innovatively define a composite loss function composed of pixel-level loss based on MS-SSIM-L1 and edge-based edge loss for LDCT denoising. In short, our approach learns a rich set of features that combine contextual information from multiple scales while maintaining the spatial details of denoised CT images. Our laboratory results indicate that compared with the existing methods, the peak signal-to-noise ratio (PSNR) value of CT images of the AAPM dataset processed by the new model is 33.6490, and the structural similarity (SSIM) value is 0.9174, which also achieves good results on the Piglet dataset with different doses. The results also show that the method removes noise and artifacts while effectively preserving CT images' architecture and grain information.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
8  
ISSN:  
0897-1889  
eISSN:  
1618-727X  
入藏号:  
WOS:000948707200001  
PubMed ID:  
36914854  
Indexed Date:  
2023-03-30  
Institution Address:  
North Univ China, Sch Informat & Commun Engn, 3 Coll Rd, Taiyuan 030051, Shanxi, Peoples R China  
Institution Address:  
North Univ China, State Key Lab Dynam Testing Technol, Taiyuan 030051, Peoples R China  
Institution Address:  
Southeast Univ, Sch Comp Sci & Engn, Nanjing 211189, Jiangsu, Peoples R China  
Institution Address:  
Southeast Univ, Jiangsu Prov Joint Int Res Lab Med Informat Proc, Nanjing 211189, Jiangsu, Peoples R China  
Institution Address:  
Southeast Univ, Key Lab Comp Network & Informat Integrat, Minist Educ, Nanjing 211189, Jiangsu, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 553. LA-Net: An End-to-End Category-Level Object Attitude Estimation Network

摘要: In category-level object pose estimation tasks, determining how to mitigate intra-class shape variations and improve pose estimation accuracy for complex objects remains a challenging problem to solve. To address this issue, this paper proposes a new network architecture, LA-Net, to efficiently ascertain object poses from features. Firstly, we extend the 3D graph convolution network architecture by introducing the LS-Layer (Linear Connection Layer), which enables the network to acquire features from different layers and perform multi-scale feature fusion. Secondly, LA-Net employs a novel attention mechanism (PSA) and a Max-Pooling layer to extract local and global geometric information, which enhances the network's ability to perceive object poses. Finally, the proposed LA-Net recovers the rotation information of an object by decoupling the rotation mechanism. The experimental results show that LA-Net can has much better accuracy in object pose estimation compared to the baseline method (HS-Pose). Especially for objects with complex shapes, its performance is 8.2% better for the 10 degrees 5 cm metric and 5% better for the 10 degrees 2 cm metric.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2079-9292  
入藏号:  
WOS:001277270400001  
Indexed Date:  
2024-08-02  
Institution Address:  
Xian Univ Sci & Technol, Sch Commun & Informat Engn, Xian 710054, Peoples R China  
Institution:

### 554. Fine-grained person-based image captioning via advanced spectrum parsing

摘要: Recent image captioning models have demonstrated remarkable performance in capturing substantial global semantic information in coarse-grained images and achieving high object coverage rates in generated captions. When applied to fine-grained images that contain heterogeneous object attributes, these models often struggle to maintain the desired granularity due to inadequate attention to local content. This paper investigates a solution for fine-grained caption generation on person-based images and heuristically proposes the Advanced Spectrum Parsing (ASP) model. Specifically, we design a novel spectrum branch to unveil the potential contour features of detected objects in the spectrum domain. We also preserve the spatial feature branch employed in existing methods, and leverage a multi-level feature extraction module to extract both spatial and spectrum features. Further more, we optimize these features, aiming to learn the spatial-spectrum correlation and complete the feature concatenation procedure via a multi-scale feature fusion module. In the inference stage, the integrated features enable the model to focus more on the local semantic regions of the person in the image. Extensive experimental results demonstrate that the proposed ASP for person-based datasets can yield promising results with both comprehensiveness and fine graininess.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001078145100001  
Indexed Date:  
2023-10-18  
Institution Address:  
Nanjing Tech Univ, Sch Comp Sci & Technol, Nanjing 211816, Peoples R China  
Institution:

### 555. Automatic Recognition of Road Damage Based on Lightweight Attentional

摘要: An efficient road damage detection system can reduce the risk of road defects to motorists and road maintenance costs to traffic management authorities, for which a lightweight end-to-end road damage detection network is proposed in this paper, aiming at fast and automatic accurate identification and classification of multiple types of road damage. The proposed technique consists of a backbone network based on a combination of lightweight feature detection modules constituted with a multi-scale feature fusion network, which is more beneficial for target identification and classification at different distances and angles than other studies. An embedded lightweight attention module was also developed that can enhance feature information by assigning weights to multi-scale convolutional kernels to improve detection accuracy with fewer parameters. The proposed model generally has higher performance and fewer parameters than other representative models. According to our practice tests, it can identify many types of road damage based on the images captured by vehicle cameras and meet the real-time detection required when piggybacking on mobile systems.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
9  
被引频次合计:  
9  
eISSN:  
1424-8220  
入藏号:  
WOS:000902775000001  
PubMed ID:  
36559968  
Indexed Date:  
2023-01-11  
Institution Address:  
Kyungpook Natl Univ, Dept Civil Engn, Daegu 37224, South Korea  
Institution:

### 556. Scale-Residual Learning Network for Scene Text Detection

摘要: Detecting incidentally captured text in the wild remains an open problem due to challenging factors including unconstrained scenarios and large scale variation. In this paper, we establish a large-scale scene text detection dataset (LS-Text), containing 36, 000 images and 270, 783 text instances with various scales and complex scenarios, to promote the research of text detection. We propose a Scale-residual Learning Network (SLN) to deal with the scale variation problem in a progressive optimization manner. Specifically, we integrate both learnable feature concatenation and feature up-sampling operator. It can effectively eliminate the residuals between the outputs of SLN and ground-truth text instances by processing both the Feature Fusion Residuals (FFR) and the Scale Transformation Residuals (STR), simultaneously. By stacking multi-scale feature maps in a deep-to-shallow manner, SLN continuously optimizes feature representation by accumulating strong semantic information and rich texture details in a scale-residual learning way. Extensive experimental results on five challenging datasets demonstrate the state-of-the-art performance of the proposed SLN model, and the challenging aspects related to real-world scenarios of the proposed LS-Text dataset. Both the source code of SLN and the LS-Text dataset are available at https://github.com/SLN-Text-Detection.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
14  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:000670583300017  
Indexed Date:  
2021-07-22  
Institution Address:  
Univ Chinese Acad Sci, Sch Comp Sci & Technol, Beijing 101408, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 101408, Peoples R China  
Institution Address:  
SUNY Albany, Comp Sci Dept, Albany, NY 12222 USA  
Institution Address:  
Chinese Acad Sci, Inst Software, Beijing 100864, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 557. Attention mechanism based multi-scale feature extraction of bearing

摘要: Effective bearing fault diagnosis is vital for the safe and reliable operation of rotating machinery. In practical applications, bearings often work at various rotational speeds as well as load conditions. Yet, the bearing fault diagnosis under multiple conditions is a new subject, which needs to be further explored. Therefore, a multi-scale deep belief network (DBN) method integrated with attention mechanism is proposed for the purpose of extracting the multi-scale core features from vibration signals, containing four primary steps: preprocessing of multi-scale data, feature extraction, feature fusion, and fault classification. The key novelties include multi-scale feature extraction using multiscale DBN algorithm, and feature fusion using attention mechanism. The benchmark dataset from University of Ottawa is applied to validate the effectiveness as well as advantages of this method. Furthermore, the aforementioned method is compared with four classical fault diagnosis methods reported in the literature, and the comparison results show that our proposed method has higher diagnostic accuracy and better robustness.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
4  
ISSN:  
1004-4132  
入藏号:  
WOS:001101657300020  
Indexed Date:  
2023-12-17  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Automat Engn, Nanjing 211106, Peoples R China  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, State Key Lab Mech & Control Mech Struct, Nanjing 211106, Peoples R China  
Institution Address:  
Nanjing Tech Univ, Coll Elect Engn & Control Sci, Nanjing 211816, Peoples R China  
Institution:  
Institution:  
Institution:

### 558. Multiscale Expression Recognition Based on Feature Selection and

摘要: In the expression recognition task, due to the diversity and uncertainty of facial features, it is easy to have problems such as missing features and low feature extraction rate in the feature extraction stage. At the same time, a large number of redundant features will be accumulated in the network training process with feature reuse structure, which will affect the feature quality. To solve the above problems, this paper proposes a residual multiscale feature fusion attentional network (RMFANet) based on feature filtering and improved convolution. Referring to the idea of blue print separable convolution and dilated convolution, the improved convolution is designed and introduced, so that the convolution can be separated more effectively and the efficiency of feature extraction can be improved. Based on the improved convolution model, a multi-scale parallel feature extraction path is designed and introduced to enrich the feature information. The feature screening module is designed and introduced to reduce the redundant features generated in the process of model training, screen out high-quality features and improve the quality of features. A shallow input feature processing layer is designed and introduced to simplify the network structure and reduce the computational complexity. Channel attention mechanism is introduced to highlight local key feature information. Finally, the SMU activation function is introduced to improve the nonlinear capability of the model. It can be seen from the experimental results that the model can achieve 70.298% and 96.566% recognition accuracy on Fer2013 data set and CK~+ data set respectively on the premise of low parameter size and calculation cost, which has better robustness than the traditional algorithm.  
摘要:  
在表情识别任务中由于人脸特征的多样性和不确定性,导致在特征提取阶段容易出现特征缺失以及特征提取率低下等问题,与此同时,在具有特征复用结构的网络训练过程中还会堆积大量冗余特征,从而影响特征质量。针对以上问题,提出了一种基于特征筛选结合改进卷积的残差多尺度特征融合注意力机制模型(residual multiscale feature fusion attentional network,RMFANet)。参考蓝图可分离卷积以及空洞卷积的思想,设计并引入了改进后的卷积形式,从而更有效地将卷积进行分离,提升特征提取效能;在改进后卷积模式的基础上设计并引入了多尺度并行特征提取通路,丰富了特征信息;设计并引入了特征筛选模块,以减少模型训练过程中产生的冗余特征,同时筛选出优质特征,提升特征质量;设计并引入了浅层输入特征处理层,以简化网络结构,降低计算复杂度;引入通道注意力机制,以突出局部关键特征信息;最后引入SMU激活函数,从而提升模型的非线性能力。通过实验结果可以看出,该模型可以在保证较低参数量以及计算成本的前提条件下在Fer2013数据集以及CK~+数据集上分别取得70.298%和96.566%的识别准确率,相比较传统算法而言具有更好的鲁棒性。  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7657633  
Indexed Date:  
2024-04-26  
Institution Address:  
北京建筑大学电气与信息工程学院, 北京 100044, 中国  
Institution Address:  
College of Electrical and Information Engineering,Beijing University of Civil Engineering and Architecture, Beijing 100044, China  
Institution:  
北京建筑大学电气与信息工程学院  
Institution:  
College of Electrical and Information Engineering,Beijing University of Civil Engineering and Architecture

### 559. YOLOv8-WTDD: multi-scale defect detection algorithm for wind turbines

摘要: In addressing the challenges of wind turbine defect detection, such as different defect scales in UAV aerial photography, interference from different lighting conditions, and small-sized target defects leading to low detection accuracy and inaccurate localization, a YOLOv8-WTBB model based on YOLOv8 is proposed. Firstly, the Diverse Branch Block is designed to enhance multi-scale feature fusion capabilities. Next, the Receptive-Field Attention Convolution is introduced to focus on the spatial features of the receptive field, increasing the distinction between target features and the surrounding environment. Finally, introducing the Minimum Point Distance Intersection over the Union bounding box regression loss function notably improves localization accuracy in object detection and accelerates model convergence. Experimental results demonstrate that the proposed algorithm significantly outperforms the baseline network, with a 4.3% improvement in mean average precision, achieving 89.1%, and a 7.4% increase in mean average recall, reaching 84.8%.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0920-8542  
eISSN:  
1573-0484  
入藏号:  
WOS:001336829400003  
Indexed Date:  
2024-10-25  
Institution Address:  
Shanghai Univ Elect Power, Dept Comp Sci & Technol, Shanghai, Peoples R China  
Institution:

### 560. Computationally Lightweight Hyperspectral Image Classification Using a

摘要: Convolutional networks have been widely used for the classification of hyperspectral images; however, such networks are notorious for their large number of trainable parameters and high computational complexity. Additionally, traditional convolution-based methods are typically implemented as a simple cascade of a number of convolutions using a single-scale convolution kernel. In contrast, a lightweight multiscale convolutional network is proposed, capitalizing on feature extraction at multiple scales in parallel branches followed by feature fusion. In this approach, 2-D depthwise convolution is used instead of conventional convolution to reduce network complexity without sacrificing classification accuracy. Furthermore, multiscale channel attention (MSCA) is also employed to selectively exploit discriminative capability across various channels. To do so, multiple 1-D convolutions with varying kernel sizes provide channel attention at multiple scales, again with the goal of minimizing network complexity. Experimental results reveal that the proposed network not only outperforms other competing lightweight classifiers in terms of classification accuracy, but also exhibits a lower number of parameters as well as significantly less computational cost.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
9  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001021284400009  
Indexed Date:  
2023-07-19  
Institution Address:  
Changan Univ, Sch Elect & Control Engn, Xian 710064, Peoples R China  
Institution Address:  
Mississippi State Univ, Dept Elect & Comp Engn, Mississippi, MS 39762 USA  
Institution:  
Institution:

### 561. Improved YOLOv5-based Underwater Infrared Garbage Detection Algorithm

摘要: An improved object detection method (YOLO with EffectiveSE, Focal-EIOU, DCNv2, CARAFE, and DyHead) is proposed based on YOLOv5 to address issues in underwater waste infrared target detection, such as blurred boundary details, low image quality, and the presence of various irregular or damaged coverings. The InceptionNeXt network is selected as the backbone network to enhance the model's expressive power and feature extraction capability. Additionally, the EffectiveSE attention mechanism is introduced in the feature fusion layer to adaptively learn the importance of feature channels and selectively weight them. Deformable convolutions are used to replace the C3 module in the original model, enabling it to better perceive the shapes and details of the targets. Moreover, the CARAFE operator is employed to replace the upsampling module, thereby enhancing the representation ability of the fine-grained features and avoiding information loss. In terms of the loss function, the Focal-EIOU loss function is adopted to improve the accuracy of the model in target localization and bounding box regression. Finally, DyHead is introduced to replace the head of YOLOv5, thereby enhancing the model accuracy via dynamic receptive field mechanisms and multiscale feature fusion. The improved EFDCD-YOLO model is applied to underwater waste infrared target detection and compared to the YOLOv5 model. The model achieves a 21.4% improvement in precision (P), 9.7% improvement in recall (R), and 13.6% improvement in mean average precision (mAP). The experimental results demonstrate that EFDCD-YOLO effectively enhances the detection performance in underwater waste infrared target detection scenarios and effectively meets the requirements of underwater infrared target detection.  
摘要:  
针对水下废弃物红外目标检测中出现的检测目标边界细节模糊、图像质量低和存在各种不规则形状或损坏的覆盖物等问题,本文提出了一种基于YOLOv5的改进目标检测方法(EFDCD-YOLO)。在主干网络中选择InceptionNeXt网络,以增强模型的表达能力和特征提取能力。其次,在特征融合层中通过加入EffectiveSE注意力机制,自适应地学习特征通道的重要性,并进行选择性加权。采用可变形卷积替代原模型中的C3模块,使模型能够更好地感知目标的形状和细节信息。此外,将CARAFE算子替代上采样模块,增强对细粒度特征的表现能力,避免信息丢失。在损失函数方面,采用Focal- EIOU损失函数,以提高模型对目标定位和边界框回归的准确性。最后,引入DyHead替换YOLOv5中的头部,通过动态感受野机制和多尺度的特征融合方式,提升模型的准确性。将改进后的EFDCDYOLO模型应用于水下废弃物红外目标检测,相比于YOLOv5模型,改进后的模型在准确率(P)、召回率(R)和平均精度(mAP)方面分别提升了21.4%、9.7%和13.6%。实验结果表明,EFDCD-YOLO能够有效地提升水下废弃物红外目标检测场景的性能,更好地满足水下废弃物红外目标检测的需求。  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1001-8891  
入藏号:  
CSCD:7806650  
Indexed Date:  
2024-11-23  
Institution Address:  
安徽工业大学计算机科学与技术学院, 马鞍山, 安徽 243032, 中国  
Institution Address:  
School of Computer Science and Technology,Anhui University of Technology, Maanshan, Anhui 243032, China  
Institution:  
安徽工业大学计算机科学与技术学院  
Institution:  
School of Computer Science and Technology,Anhui University of Technology

### 562. High-Precision Segmentation of Buildings with Small Sample Sizes Based

摘要: In order to improve the accuracy of the segmentation of buildings with small sample sizes, this paper proposes a building-segmentation network, ResFAUnet, with transfer learning and multi-scale feature fusion. The network is based on AttentionUnet. The backbone of the encoder is replaced by the ResNeXt101 network for feature extraction, and the attention mechanism of the skip connection is preserved to fuse the shallow features of the encoding part and the deep features of the decoding part. In the decoder, the feature-pyramid structure is used to fuse the feature maps of different scales. More features can be extracted from limited image samples. The proposed network is compared with current classical semantic segmentation networks, Unet, SuUnet, FCN, and SegNet. The experimental results show that in the dataset selected in this paper, the precision indicators of ResFAUnet are improved by 4.77%, 2.3%, 2.11%, and 1.57%, respectively, compared with the four comparison networks.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
eISSN:  
2072-4292  
入藏号:  
WOS:000988113700001  
Indexed Date:  
2023-06-05  
Institution Address:  
Hohai Univ, Coll Mech & Elect Engn, Changzhou 213022, Peoples R China  
Institution Address:  
Hohai Univ, Jiangsu Key Lab Special Robot Technol, Changzhou 213022, Peoples R China  
Institution:  
Institution:

### 563. Wavelet-Based, Blur-Aware Decoupled Network for Video Deblurring

摘要: Video deblurring faces a fundamental challenge, as blur degradation comprehensively affects frames by not only causing detail loss but also severely distorting structural information. This dual degradation across low- and high-frequency domains makes it challenging for existing methods to simultaneously restore both structural and detailed information through a unified approach. To address this issue, we propose a wavelet-based, blur-aware decoupled network (WBDNet) that innovatively decouples structure reconstruction from detail enhancement. Our method decomposes features into multiple frequency bands and employs specialized restoration strategies for different frequency domains. In the low-frequency domain, we construct a multi-scale feature pyramid with optical flow alignment. This enables accurate structure reconstruction through bottom-up progressive feature fusion. For high-frequency components, we combine deformable convolution with a blur-aware attention mechanism. This allows us to precisely extract and merge sharp details from multiple frames. Extensive experiments on benchmark datasets demonstrate the superior performance of our method, particularly in preserving structural integrity and detail fidelity.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001418441400001  
Indexed Date:  
2025-02-17  
Institution Address:  
Mahasarakham Univ, Fac Informat, Dept Comp Sci, Maha Sarakham 44150, Thailand  
Institution:

### 564. Multimodal and multiscale feature fusion for weakly supervised video

摘要: Weakly supervised video anomaly detection aims to detect anomalous events with only video-level labels. In the absence of boundary information for anomaly segments, most existing methods rely on multiple instance learning. In these approaches, the predictions for unlabeled video snippets are guided by the classification of labeled untrimmed videos. However, these methods do not account for issues such as video blur and visual occlusion, which can hinder accurate anomaly detection. To address these issues, we propose a novel weakly supervised video anomaly detection method that fuses multimodal and multiscale features. Firstly, RGB and optical flow snippets are input into pre-trained I3D to extract appearance and motion features. Then, we introduce an Attention De-redundancy (AD) module, which employs an attention mechanism to filter out task-irrelevant redundancy in these appearance and motion features. Next, to mitigate the effects of video blurring and visual occlusion, we propose a Multi-scale Feature Learning module. This module captures long-term and short-term temporal dependencies among video snippets to provide global and local guidance for blurred or occluded video snippets. Finally, to effectively utilize the discriminative features of different modalities, we propose an Adaptive Feature Fusion module. This module adaptively fuses appearance and motion features based on their respective feature weights. Extensive experimental results demonstrate that our proposed method outperforms mainstream unsupervised and weakly supervised methods in terms of AUC. Specifically, our proposed method achieves 97.00% AUC and 85.31% AUC on two benchmark datasets, i.e., ShanghaiTech and UCF-Crime, respectively.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2045-2322  
入藏号:  
WOS:001328801300054  
PubMed ID:  
39354033  
Indexed Date:  
2024-10-16  
Institution Address:  
Beijing Informat Sci & Technol Univ, Key Lab, Minist Educ Optoelect Measurement Technol & Instru, Beijing 100192, Peoples R China  
Institution Address:  
Beijing Informat Sci & Technol Univ, Key Lab Informat & Commun Syst, Minist Informat Ind, Beijing 100192, Peoples R China  
Institution Address:  
Beijing Informat Sci & Technol Univ, Sch Instrument Sci & Optoelect Engn, Beijing 100192, Peoples R China  
Institution:  
Institution:  
Institution:

### 565. A Study of Occluded Person Re-Identification for Shared Feature Fusion

摘要: The human body is often occluded by a variety of obstacles in the monitoring system, so occluded person re-identification is still a long-standing challenge. Recent methods based on pose guidance or external semantic clues have improved the representation and related performance of features; there are still problems, such as weak model representation and unreliable semantic clues. To solve the above problems, we proposed a feature extraction network, named shared feature fusion with pose-guided and unsupervised semantic segmentation (SFPUS). This network will extract more discriminative features and reduce the occlusion noise on pedestrian matching. Firstly, the multibranch joint feature extraction module (MFE) is used to extract feature sets containing pose information and high-order semantic information. This module not only provides robust extraction capabilities but can also precisely segment occlusion and the body. Secondly, in order to obtain multiscale discriminant features, the multiscale correlation feature matching fusion module (MCF) is used to match the two feature sets, and the Pose-Semantic Fusion Loss is designed to calculate the similarity of the feature sets between different modes and fuse them into a feature set. Thirdly, to solve the problem of image occlusion, we use unsupervised cascade clustering to better prevent occlusion interference. Finally, performances of the proposed method and various existing methods are compared on the Occluded-Duke, Occluded-ReID, Market-1501 and Duke-MTMC datasets. The accuracy of Rank-1 reached 65.7%, 80.8%, 94.8% and 89.6%, respectively, and the mAP accuracy reached 58.8%, 72.5%, 91.8% and 80.1%. The experiment results demonstrate that our proposed SFPUS holds promising prospects and performs admirably compared with state-of-the-art methods.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001364304300001  
Indexed Date:  
2024-12-03  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Automat, Xian Key Lab Adv Control & Intelligent Proc, Xian 710000, Peoples R China  
Institution:

### 566. Asymptotic Feature Pyramid Network for Labeling Pixels and Regions

摘要: Multi-scale features are crucial in encoding objects with varying scales in vision tasks. The classic top-down and bottom-up feature pyramid networks are a common strategy for multi-scale feature extraction. However, these approaches suffer from the loss or degradation of feature information, which impairs the fusion effect of non-adjacent levels. In this paper, we propose an Asymptotic Feature Pyramid Network (AFPN) that supports direct interaction between non-adjacent levels. AFPN starts by fusing two adjacent low-level features and asymptotic incorporates higher-level features into the fusion process. This fusion way avoids the significant semantic gap between non-adjacent levels. Adaptive spatial fusion operation is further used to mitigate potential multi-object information conflicts during feature fusion at each spatial location. To reduce parameters, computational requirements, and inference speed, we propose a Lightweight Asymptotic Feature Pyramid Network (LightAFPN) that uses the concept of reparametrization. We evaluate the proposed method on the MS-COCO 2017, PASCAL VOC and Cityscapes datasets in both object detection and semantic segmentation frameworks. Experimental evaluation shows that our method achieves more competitive results than other state-of-the-art feature pyramid networks. The code is available at https://github.com/gyyang23/AFPN.  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:001409508700002  
Indexed Date:  
2025-02-07  
Institution Address:  
Zhejiang Univ Technol, Coll Comp Sci & Technol, Hangzhou 310023, Peoples R China  
Institution Address:  
Zhejiang Univ, Coll Comp Sci & Technol, Hangzhou 310027, Peoples R China  
Institution:  
Institution:

### 567. A multi-scale attributes fusion model for travel mode identification

摘要: Travel mode recognition is a key issue in urban planning and transportation research. While traditional travel surveys use manual data collection and have limited coverage, poor timeliness, and insufficient sample capacity, recent advancements in Global Positioning System (GPS) technology allow large-scale data collection and offer novel opportunities to enhance travel mode recognition. However, existing studies often neglect regular differences and changes in motion states across different travel modes and fail to fully integrate multi-scale spatio-temporal features, which limits the accurate classification of travel modes. To fill this gap, this study proposes a multi-scale spatio-temporal attribute fusion (MSAF) model for precise travel mode identification using solely GPS trajectories without altering their sampling rate. The MSAF model segments GPS trajectories into various temporal and spatial scales, extracting local motion states and spatial features at multiple scales. The spatio-temporal feature extraction module is constructed to extract local motion states and capture spatio-temporal dependencies. Additionally, the model incorporates a multi-scale feature fusion module, which effectively combines features of various scales through a series of fusion techniques to obtain a comprehensive representation, enabling automatic and accurate travel mode identification. Experiments on real-world datasets, including the GeoLife Trajectories dataset and the Sussex-Huawei Locomotion-Transportation (SHL) dataset, demonstrate the effectiveness of the MSAF model, achieving a competitive accuracy of 95.16% and 91.70%. This represents an improvement of 2.50% to 7.95% and 0.8% to 6.62% over several state-of-the-art baselines, effectively addressing sample imbalance challenges. Moreover, the experiments demonstrate the significant role of multiscale feature fusion in improving model performance.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1009-5020  
eISSN:  
1993-5153  
入藏号:  
WOS:001286766300001  
Indexed Date:  
2024-08-14  
Institution Address:  
Fuzhou Univ, Acad Digital China Fujian, Fuzhou, Peoples R China  
Institution Address:  
Fuzhou Univ, Lab Spatial Data Min & Informat Sharing, Minist Educ, Fuzhou, Peoples R China  
Institution:  
Institution:

### 568. Lightweight detection method of coal gangue based on multispectral and

摘要: Aiming at the problems of poor identification effect and low separation efficiency of existing coal gangue identification and separation methods, this paper puts forward a lighter coal gangue identification and detection scheme based on YOLOv5s (You Only Look Once Version-5s) and multispectral image technology. Set up a multispectral acquisition system and shoot the required data sets. Three bands with high recognition accuracy and low correlation are selected from 25 bands in the spectrum to form RGB images for model detection. A depth separable convolution is introduced into the feature fusion network, which makes the model lighter, better fuses multiscale feature information, adds attention mechanism to the network, strengthens the objective attention, improves the ability of dense object detection and antibackground interference, reduces the complexity of the model, and improves the detection accuracy. The experimental results show that the improved YOLOv5s-MobileNet model has a mean average precision (mAP) of 98.88% on the coal and gangue test set. Compared with YOLOv5s algorithm, the model size is reduced by 12.5%, and the mAP is increased by 12.88%, which is beneficial to edge deployment to the greatest extent.  
Web of Science 核心合集中的 "被引频次":  
4  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1939-2699  
eISSN:  
1939-2702  
入藏号:  
WOS:000956377700001  
Indexed Date:  
2023-04-10  
Institution Address:  
Anhui Univ Sci & Technol, State Key Lab Min Response & Disaster Prevent & Co, Huainan, Peoples R China  
Institution Address:  
Anhui Univ Sci & Technol Huainan, Sch Elect & Informat Engn, Huainan, Anhui, Peoples R China  
Institution Address:  
Anhui Univ Sci & Technol, Collaborat Innovat Ctr Mine Intelligent Equipment, Huainan, Peoples R China  
Institution Address:  
Anhui Univ Sci & Technol Huainan, Sch Artificial Intelligence, Huainan, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 569. A Deep Motion Deblurring Network Using Channel Adaptive Residual Module

摘要: In this paper, we solve the problem of dynamic scenes deblurring with motion blur. Restoration of images in the presence of motion blur necessitates a network design that the receptive field can completely cover all areas that need to be deblurred, while the existing network increases the receptive field by continuously stacking the ordinary convolutional layer or increasing the size of the convolution kernel. However, these methods inevitably increase the computational burden of the network. We propose a novel architecture consisting of a channel adaptive residual module. Different features of the blurred image are extracted and distributed on each feature channel. Our network can calculate the weight of each channel through learning, and extract the image features adaptively according to different degrees of blurring and importance of information. We embed the module in a modified encoder-decoder design with skip connections to achieve multi-scale feature fusion for further performance improvement. The extensive comparison with the existing techniques in the baseline dynamic scene deblurring dataset shows that the proposed network can effectively realize image deblurring, and the accuracy and speed are comparable with the existing techniques.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
2169-3536  
入藏号:  
WOS:000647306700001  
Indexed Date:  
2021-05-13  
Institution Address:  
Hangzhou Dianzi Univ, Sch Elect & Informat, Hangzhou 310018, Peoples R China  
Institution Address:  
Hangzhou Dianzi Univ, Zhejiang Prov Key Lab Equipment Elect, Hangzhou 310018, Peoples R China  
Institution:  
Institution:

### 570. Enhancing small object detection in point clouds with self-attention

摘要: The development of point cloud-based object detection in the field of autonomous driving has been rapid. However, it is undeniable that the issue of detecting small objects with high precision remains an urgent challenge. To address this issue, we introduce a single-stage 3D detection network, termed self-attention voting-single stage detection (SAV-SSD). It directly extracts feature information from the raw point cloud data and introduces an innovative self-attention voting mechanism to generate center points through weighted voting based on feature correlations. Compared with the feature prediction, we make an additional prediction of the center point, which can better control the position and size of the bounding boxes to improve the accuracy and stability of the predictions. To capture more features of small objects, cross multi-scale feature fusion is designed to establish connections between deep and shallow features. Experimental results demonstrate that SAV-SSD significantly improves the accuracy of pedestrian and cyclist detection while maintaining real-time performance. On the KITTI dataset, SAV-SSD outperforms many state-of-the-art 3D object detection methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0091-3286  
eISSN:  
1560-2303  
入藏号:  
WOS:001232239600029  
Indexed Date:  
2024-06-13  
Institution Address:  
Hubei Univ Technol, Sch Elect & Elect Engn, Wuhan, Peoples R China  
Institution Address:  
Tianjin Univ Sci & Technol, Coll Artificial Intelligence, Tianjin, Peoples R China  
Institution:  
Institution:

### 571. A rotatable battery recognition method based on improved YOLOv5

摘要: To realise end-to-end visual identification, positioning, and angle detection of cylindrical batteries, a rotated object recognition method based on YOLOv5 is proposed. Firstly, aiming at the problems of battery appearance scale variation and surface reflection, a recursive gated convolution and feature fusion module was added to the neck network to enhance the multi-scale feature extraction. Secondly, considering the boundary problem of angle range, a circular smooth label was introduced after the prediction network, and the logistic regression cross-entropy was used to realise rotation angle classification. Finally, a SIoU intersection ratio model was used to introduce an angle vector penalty index. The experimental results show that the parameters of rotated object detection model are reasonably optimised on the cylindrical battery dataset. The model accuracy reaches 98.6%, the recall rate reaches 97.2%, and the inference speed of single frame image reaches 10.5 ms, which meets the performance requirements of practical applications.  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1748-1279  
eISSN:  
1748-1287  
入藏号:  
WOS:001238114800005  
Indexed Date:  
2024-06-09  
Institution Address:  
Ningbo Polytech, Sch Elect & Informat Engn, Ningbo 315800, Zhejiang, Peoples R China  
Institution Address:  
Ningbo Univ, Fac Mech Engn & Mech, Ningbo 315210, Zhejiang, Peoples R China  
Institution:  
Institution:

### 572. Convolutional Neural Network Defect Detection Algorithm for Wire Bonding

摘要: To address the challenges of complex backgrounds, small defect sizes, and diverse defect types in defect detection of wire bonding X-ray images, this paper proposes a convolutional-neural-network-based defect detection method called YOLO-CSS. This method designs a novel feature extraction network that effectively captures semantic features from different gradient information. It utilizes a self-adaptive weighted multi-scale feature fusion module called SMA which adaptively weights the contribution of detection results based on different scales of feature maps. Simultaneously, skip connections are employed at the bottleneck of the network to ensure the integrity of feature information. Experimental results demonstrate that on the wire bonding X-ray defect image dataset, the proposed algorithm achieves mAP 0.5 and mAP 0.5-0.95 values of 97.3% and 72.1%, respectively, surpassing the YOLO series algorithms. It also exhibits certain advantages in terms of model size and detection speed, effectively balancing detection accuracy and speed.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2072-666X  
入藏号:  
WOS:001073742800001  
PubMed ID:  
37763900  
Indexed Date:  
2023-10-12  
Institution Address:  
State Key Lab Precis Elect Mfg Technol & Equipment, Guangzhou 510006, Peoples R China  
Institution Address:  
Guangdong Univ Technol, Sch Electromech Engn, Guangzhou 510006, Peoples R China  
Institution:  
State Key Lab Precis Elect Mfg Technol & Equipment  
Institution:

### 573. Single Image Deraining Integrating Physics Model and Density-Oriented

摘要: Although advanced single image deraining methods have been proposed, their generalization ability to real-world images is usually limited, especially when dealing with rain patterns of different densities, shapes, and directions. In order to improve the robustness and generalization of these deraining methods, we propose a novel density-aware single image deraining method with gated multi-scale feature fusion, which consists of two stages. In the first stage, a sophisticated physics model is leveraged for initial deraining and a network branch is utilized for rain density estimation to guide the subsequent refinement. The second stage of model-independent refinement is realized using conditional Generative Adversarial Network (cGAN), attempting to eliminate artifacts and improve the restoration quality. Extensive experiments have been conducted on the representative synthetic rain datasets and real rain scenes, demonstrating the superiority of our method in terms of effectiveness and generalization ability, which outperforms the state-of-the-arts.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
1070-9908  
eISSN:  
1558-2361  
入藏号:  
WOS:000690436100001  
Indexed Date:  
2021-09-06  
Institution Address:  
Wuhan Univ, Sch Elect Informat, Wuhan 430072, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Remote Sensing & Informat Engn, Wuhan 430072, Peoples R China  
Institution Address:  
Natl Univ Singapore, Singapore Inst Neurotechnol, Singapore 117456, Singapore  
Institution Address:  
Peng Cheng Lab, Shenzhen 518055, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 574. ATT-YOLOv5-Ghost: water surface object detection in complex scenes

摘要: In recent years, Unmanned Surface Vehicles (USVs) have been widely used in water surface monitoring and management. The main problems of the USVs-based water surface object detection method are that the features will be lost when downsampling complex water surface environment images, resulting in low detection accuracy. Moreover, the number of parameters and calculation amount of these models are too much, which will seriously affect the speed of training and detection. Therefore, this paper proposed the ATT-YOLOv5-Ghost algorithm. First, we added an Efficient Channel Attention (ECA) module to each CSP1 unit of the backbone CSPDarknet, which solves the problem of accuracy drop caused by multi-scale feature loss during downsampling. Second, we proposed a method combining ECA and Ghost modules. In the process of feature fusion, the problems such as the increase of parameters, slow detection speed and repeated gradient calculation caused by too complex algorithm were solved. The ATT-YOLOv5-Ghost algorithm improves the detection accuracy by 4.6% compared with the baseline. The FPS can reach 64.9, and the computational amount is reduced by 8.9%. The algorithm complexity was significantly reduced.  
Web of Science 核心合集中的 "被引频次":  
10  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001050358900001  
Indexed Date:  
2023-08-28  
Institution Address:  
Harbin Univ Sci & Technol, Sch Automat, Heilongjiang Prov Key Lab Complex Intelligent Syst, Harbin 150080, Peoples R China  
Institution:

### 575. No-Reference Image Quality Assessment Based on a Multitask Image

摘要: When image quality is evaluated, the human visual system (HVS) infers the details in the image through its internal generative mechanism. In this process, the HVS integrates both local and global information about the image, utilizes contextual information to restore the original image information, and compares it with the distorted image information for image quality evaluation. Inspired by this mechanism, a no-reference image quality assessment method is proposed based on a multitask image restoration network. The multitask image restoration network generates a pseudo-reference image as the main task and produces a structural similarity index measure map as an auxiliary task. By mutually promoting the two tasks, a higher-quality pseudo-reference image is generated. In addition, when predicting the image quality score, both the quality restoration features and the difference features between the distorted and reference images are used, thereby fully utilizing the information from the pseudo-reference image. In order to facilitate the model's ability to extract both global and local features, we introduce a multi-scale feature fusion module. Experimental results demonstrate that the proposed method achieves excellent performance on both synthetically and authentically distorted databases.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
eISSN:  
2076-3417  
入藏号:  
WOS:001003863200001  
Indexed Date:  
2023-06-20  
Institution Address:  
Shenzhen Univ, Dept Artificial Intelligence, Shenzhen 518060, Peoples R China  
Institution Address:  
Educ Univ Hong Kong, Dept Math & Informat Technol, Hong Kong, Peoples R China  
Institution Address:  
Univ Massachusetts Lowell, Dept Elect & Comp Engn, Lowell, MA 01854 USA  
Institution:  
Institution:  
Institution:

### 576. Salient feature network for semantic segmentation

摘要: In the encoding stage, some existing semantic segmentation networks capture rich multi-scale context information. However, multi-scale approaches do not pay attention to the correlation between different scale feature maps in the multi-scale feature fusion stage. In the decoding stage, simple fusion of high- and low-dimensional channel is used to improve the semantic segmentation, but simple fusion suffers from the defect that the segmentation boundary is not sufficiently clear. In this paper, a salient feature network is proposed to address these two disadvantages. For the first shortcoming, an atrous spatial pyramid pooling with Euclidean distance similarity (EDS-ASPP) module is proposed to enhance the representation of high-level semantic information features, that is, to boost meaningful features, while suppressing weak ones. Therefore, this module can solve the segmentation error inside objects. For the second deficiency, a supplementary details (SD) module is proposed to rearrange the low-level spatial details and the activation graph obtained from the EDS-ASPP module in the decoding stage. The function of this module is to repair the edge details lost during the downsampling process. The proposed model achieves a 73.45% mIoU on PASCAL VOC2012 and a 64.27% mIoU on Cityscapes.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:000696425200001  
Indexed Date:  
2021-10-12  
Institution Address:  
Xiangtan Univ, Sch Automat & Elect Informat, Xiangtan 411105, Peoples R China  
Institution Address:  
Xiangtan Univ, Sch Math & Computat Sci, Xiangtan 411105, Peoples R China  
Institution Address:  
Shanghai Aerosp Control Technol Inst, Shanghai 201109, Peoples R China  
Institution:  
Institution:  
Institution:  
Shanghai Aerosp Control Technol Inst

### 577. Integrating Detailed Features and Global Contexts for Semantic

摘要: Semantic segmentation of ultrahigh-resolution (UHR) remote sensing images is a fundamental task for many downstream applications. Achieving precise pixel-level classification is paramount for obtaining exceptional segmentation results. This challenge becomes even more complex due to the need to address intricate segmentation boundaries and accurately delineate small objects within the remote sensing imagery. To meet these demands effectively, it is critical to integrate two crucial components: global contextual information and spatial detail feature information. In response to this imperative, the multilevel context-aware segmentation network (MCSNet) emerges as a promising solution. MCSNet is engineered to not only model the overarching global context but also extract intricate spatial detail features, thereby optimizing segmentation outcomes. The strength of MCSNet lies in its two pivotal modules, the spatial detail feature extraction (SDFE) module and the refined multiscale feature fusion (RMFF) module. Moreover, to further harness the potential of MCSNet, a multitask learning approach is employed. This approach integrates boundary detection and semantic segmentation, ensuring that the network is well-rounded in its segmentation capabilities. The efficacy of MCSNet is rigorously demonstrated through comprehensive experiments conducted on two established international society for photogrammetry and remote sensing (ISPRS) 2-D semantic labeling datasets: Potsdam and Vaihingen. These experiments unequivocally establish MCSNet stands as a pioneering solution, that delivers state-of-the-art performance, as evidenced by its outstanding mean intersection over union (mIoU) and mean $F1$ -score (mF1) metrics. The code is available at: https://github.com/WUTCM-Lab/MCSNet.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
被引频次合计:  
6  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001219598000016  
Indexed Date:  
2024-05-23  
Institution Address:  
Wuhan Univ Technol, Sch Comp Sci & Artificial Intelligence, Wuhan 430070, Peoples R China  
Institution Address:  
Wuhan Univ Technol, Sanya Sci & Educ Innovat Pk, Sanya 572000, Peoples R China  
Institution Address:  
Shanghai Artificial Intelligence Lab, Shanghai 200232, Peoples R China  
Institution Address:  
Wuhan Univ Technol, Chongqing Res Inst, Chongqing 401122, Peoples R China  
Institution Address:  
Wuhan Huaxia Inst Technol, Sch Informat Engn, Wuhan 430223, Peoples R China  
Institution Address:  
Qiongtai Normal Univ, Sch Informat Sci & Technol, Haikou 571127, Peoples R China  
Institution Address:  
Fuzhou Univ, Coll Phys & Informat Engn, Fuzhou 350108, Peoples R China  
Institution Address:  
Tech Univ Munich, Data Sci Earth Observat, D-80333 Munich, Germany  
Institution:  
Institution:  
Institution:  
Shanghai Artificial Intelligence Lab  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 578. Multi-Scale Fusion Model Based on Gated Recurrent Unit for Enhancing

摘要: Accurate prediction of the state-of-charge (SOC) of battery energy storage system (BESS) is critical for its safety and lifespan in electric vehicles. To overcome the imbalance of existing methods between multi-scale feature fusion and global feature extraction, this paper introduces a novel multi-scale fusion (MSF) model based on gated recurrent unit (GRU), which is specifically designed for complex multi-step SOC prediction in practical BESSs. Pearson correlation analysis is first employed to identify SOC-related parameters. These parameters are then input into a multi-layer GRU for point-wise feature extraction. Concurrently, the parameters undergo patching before entering a dual-stage multi-layer GRU, thus enabling the model to capture nuanced information across varying time intervals. Ultimately, by means of adaptive weight fusion and a fully connected network, multi-step SOC predictions are rendered. Following extensive validation over multiple days, it is illustrated that the proposed model achieves an absolute error of less than 1.5% in real-time SOC prediction.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
2196-5625  
eISSN:  
2196-5420  
入藏号:  
WOS:001194022300012  
Indexed Date:  
2024-04-18  
Institution Address:  
Univ Sci & Technol Beijing, Sch Comp & Commun Engn, Beijing, Peoples R China  
Institution Address:  
Univ Sci & Technol Beijing, Sch Mech Engn, Beijing, Peoples R China  
Institution:  
Institution:

### 579. Efficient Parallel Branch Network With Multi-Scale Feature Fusion for

摘要: Image-based segmentation of overhead power lines is critical for power line inspection. Real-time segmentation helps the inspection robot avoid obstacles or land on the wire during the inspection task. It is challenging for several studies to achieve real-time overhead power line segmentation with high accuracy. In addition, cluttered background brings great difficulties to overhead power lines segmentation. To address these issues, an efficient parallel branch network for real-time overhead power line segmentation is proposed. Our framework combines a context branch that generates useful global information with a spatial branch that preserves high-resolution segmentation details. The asymmetric factorized depth-wise bottleneck (AFDB) module is designed in the context branch to achieve more efficient short-range feature extraction and provide a large receptive field. Furthermore, the subnetwork-level skip connections in the classifier are proposed to fuse long-range features and lead to high accuracy. Experiments demonstrate that our framework achieves more than 90% segmentation accuracy.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
27  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
30  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:000642012400099  
Indexed Date:  
2021-05-23  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Management & Control Complex Syst, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Artificial Intelligence, Beijing 100049, Peoples R China  
Institution Address:  
State Grid Shandong Elect Power Co, Jinan 250001, Peoples R China  
Institution:  
Institution:  
Institution:

### 580. Underwater Object Detection Based on Improved EfficientDet

摘要: Intelligent detection of marine organism plays an important part in the marine economy, and it is significant to detect marine organisms quickly and accurately in a complex marine environment for the intelligence of marine equipment. The existing object detection models do not work well underwater. This paper improves the structure of EfficientDet detector and proposes the EfficientDet-Revised (EDR), which is a new marine organism object detection model. Specifically, the MBConvBlock is reconstructed by adding the Channel Shuffle module to enable the exchange of information between the channels of the feature layer. The fully connected layer of the attention module is removed and convolution is used to cut down the amount of network parameters. The Enhanced Feature Extraction module is constructed for multi-scale feature fusion to enhance the feature extraction ability of the network to different objects. The results of experiments demonstrate that the mean average precision (mAP) of the proposed method reaches 91.67% and 92.81% on the URPC dataset and the Kaggle dataset, respectively, which is better than other object detection models. At the same time, the processing speed reaches 37.5 frame per second (FPS) on the URPC dataset, which can meet the real-time requirements. It can provide a useful reference for underwater robots to perform tasks such as intelligent grasping.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
31  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
31  
eISSN:  
2072-4292  
入藏号:  
WOS:000856678200001  
Indexed Date:  
2022-09-28  
Institution Address:  
Ocean Univ China, Coll Elect Engn, Qingdao 266100, Peoples R China  
Institution Address:  
Ocean Univ China, Sanya Oceanog Inst, Sanya 572024, Peoples R China  
Institution Address:  
Qingdao Univ Sci & Technol, Coll Automat & Elect Engn, Qingdao 266061, Peoples R China  
Institution:  
Institution:  
Institution:

### 581. B2MFuse: A Bi-Branch Multiscale Infrared and Visible Image Fusion

摘要: Infrared and visible image fusion (IVIF) is a key technique for synthesizing a comprehensive representation of a scene by exploiting diverse perceptual information. However, existing fusion methods encounter challenges in simultaneously preserving intricate texture details and extracting high-level semantic information-essential for downstream vision tasks. To address these issues, this article presents bi-branch multiscale infrared and visible image fusion network (B2MFuse), a novel bi-branch multiscale IVIF network based on joint semantics injection. The bi-branch consists of an interactive detail branch and a parallel semantic branch, both featuring dual paths for infrared and visible modality. The former employs a channel exchange strategy that maximizes the capture of modality-specific details while obtaining complementary features from the alternate modality. The latter efficiently captures semantic information and provides flexible scene knowledge guidance to the interactive detail feature extraction branch (ID-branch), facilitating the subsequent top-to-bottom multiscale feature fusion and reconstruction process. A spatial weighted channel attention fusion module (SWCAFM) is then meticulously designed to enhance the integration of crucial fine-grained features across different scales. Furthermore, a scene-perception loss function is tailored to account for variations in the original image content. The synergy between B2MFuse's advanced architecture and loss function ensures robust and superior fusion results in diverse environments, in particular enhancing human visual observation and supporting downstream visual tasks. Extensive evaluations on four public datasets demonstrate the superiority of our B2MFuse, compared with the state-of-the-art (SOTA) IVIF methods. The source code is available at: https://github.com/arkymeng/B2MFuse.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001346129100015  
Indexed Date:  
2024-11-14  
Institution Address:  
Chinese Acad Sci, Shanghai Inst Microsyst & Informat Technol, Sci & Technol Microsyst Lab, Shanghai, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun, Beijing 100049, Peoples R China  
Institution Address:  
Chinese Acad Sci, Shanghai Inst Microsyst & Informat Technol, Shanghai, Peoples R China  
Institution:  
Institution:  
Institution:

### 582. Research and Implementation of an Embedded Traffic Sign Detection Model

摘要: This study proposes an embedded traffic sign detection system, YOLOV5-MCBS, based on an enhanced YOLOv5 algorithm. This system aims to mitigate the impact of traditional target detection algorithms' high computational complexity and low detection accuracy on traffic sign detection performance, thereby improving accuracy and real-time performance. Our primary objective is to develop a lightweight network that enhances detection accuracy, enabling real-time detection on embedded systems. First, to minimize computation and model size, we replaced the original YOLOv5 algorithm's backbone feature network with a lightweight MobileNetV3 network. Subsequently, we introduced the convolutional block attention module into the neck network to optimize the feature fusion stage's attention and enhance model detection accuracy. Concurrently, we employed the bidirectional feature pyramid network in the neck layer for multi-scale feature fusion. Additionally, we incorporated a small target detection layer into the original network output layer to enhance detection performance. What's more, we transplanted the enhanced algorithm into a Raspberry Pi embedded system to validate its real-time detection performance. Finally, we conducted computer simulations to assess our algorithm's performance by comparing it with existing target detection algorithms. Experimental results suggest that the enhanced algorithm achieves an average precision mean (mAP @ 0.5) value of 95.3% and frames per second value of 91.1 on the embedded system.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1229-9138  
eISSN:  
1976-3832  
入藏号:  
WOS:001204768100003  
Indexed Date:  
2024-04-29  
Institution Address:  
Nanjing Forestry Univ, Coll Informat Sci & Technol, Nanjing 210037, Peoples R China  
Institution:

### 583. Automatic extraction and measurement of ultrasonic muscle morphological

摘要: Background: Estimating skeletal muscle force output and structure requires measurement of morphological parameters including muscle thickness, pennation angle, and fascicle length. The identification of aponeurosis and muscle fascicles from medical images is required to measure these parameters accurately. Methods: This paper introduces a multi-stage fusion and segmentation model (named MSF-Net), to precisely extract muscle aponeurosis and fascicles from ultrasound images. The segmentation process is divided into three stages of feature fusion modules. A prior feature fusion module (PFFM) is designed in the first stage to fuse prior features, thus enabling the network to focus on the region of interest and eliminate image noise. The second stage involves the addition of multi-scale feature fusion module (MS-FFM) for effective fusion of elemental information gathered from different scales. This process enables the precise extraction of muscle fascicles of varied sizes. Finally, the high-low-level feature fusion attention module (H-LFFAM) is created in the third stage to selectively reinforce features containing useful information. Results: Our proposed MSF-Net outperforms other methods and achieves the highest evaluation metrics. In addition, MSF-Net can obtain similar results to manual measurements by clinical experts. The mean deviation of muscle thickness and fascicle length was 0.18 mm and 1.71 mm, and the mean deviation of pennation angle was 0.31 degrees. Conclusions: MSF-Net can accurately extract muscle morphological parameters, which enables medical experts to evaluate muscle morphology and function, and guide rehabilitation training. Therefore, MSF-Net provides a complementary imaging tool for clinical assessment of muscle structure and function.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0041-624X  
eISSN:  
1874-9968  
入藏号:  
WOS:001101547000001  
PubMed ID:  
37883820  
Indexed Date:  
2023-12-17  
Institution Address:  
Northwestern Polytech Univ, Inst Med Res, Xian, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Sch Artificial Intelligence Opt & Elect iOPEN, Xian, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Affiliated Hosp 1, Dept Rehabil Med, Guangzhou, Peoples R China  
Institution Address:  
Xi An Childrens Hosp, Xian, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Xi An Childrens Hosp

### 584. Small object detection in UAV aerial images based on inverted residual

摘要: Aiming at the problems of complex background and too many small-size targets in UAV aerial images, a small target detection algorithm based on inverted residual attention is proposed. Firstly, an inverted residual module and an inverted residual attention module are embedded into the backbone network, while rich spatial information and deep semantic information of small targets are obtained by feature information mapping from low dimension to high dimension, thus improving the accuracy of small target detection; Secondly, in feature fusion, a multi-scale feature fusion module is established to fuse the shallow spatial information and deep semantic information, and to generate four detection heads with different sensory fields, which improves the recognition of small-size targets and reduces missed detection of small targets; Finally, a mosaic mixed data enhancement method is designed to establish the linear relationship between the data, increase the complexity of the image background and improve the robustness of the algorithm. The experimental results on data set VisDrone show that the mean average precision of this algorithm is 1.2% higher than that of DSHNet, which means that the proposed algorithm could effectively reduce missed detection and false detection of small targets in UAV aerial images.  
摘要:  
针对无人机航拍图像背景复杂、小尺寸目标较多等问题,提出了一种基于倒置残差注意力的无人机航拍图像小目标检测算法。在主干网络部分嵌入倒置残差模块与倒置残差注意力模块,利用低维向高维的特征信息映射,获得丰富的小目标空间信息和深层语义信息,提升小目标的检测精度;在特征融合部分设计多尺度特征融合模块,融合浅层空间信息和深层语义信息,并生成4个不同感受野的检测头,提升模型对小尺寸目标的识别能力,减少小目标的漏检;设计马赛克混合数据增强方法,建立数据之间的线性关系,增加图像背景复杂度,提升算法的鲁棒性。在VisDrone数据集上的实验结果表明:所提模型的平均精度均值比DSHNet模型提升了1.2%,有效改善了无人机航拍图像小目标漏检、误检的问题。  
在中国科学引文数据库中的被引频次:  
4  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1001-5965  
入藏号:  
CSCD:7445192  
Indexed Date:  
2023-07-21  
Institution Address:  
天津城建大学计算机与信息工程学院, 天津 300384, 中国  
Institution Address:  
School of Computer and Information Engineering,Tianjin Chengjian University, Tianjin 300384, China  
Institution:  
天津城建大学计算机与信息工程学院  
Institution:  
School of Computer and Information Engineering,Tianjin Chengjian University

### 585. MSFANet: A Light Weight Object Detector Based on Context Aggregation and

摘要: Accurate and reliable object detection is a fundamental component of perception system for autonomous driving. Specially, in some circumstances like autonomous driving in surface mine, there is a fact that the particularity of scene brings tremendous challenges for object detection with a series of problems caused by the multi-scale and camouflaged objects. In this paper, a multi-scale feature fusion and attention based multi-branches framework was proposed to improve the performance of object detection for above problems called MSFANet. In the proposed MSFANet, a multi-scale feature fusion module, which was used to capture the rich context features for multi-scale high level feature maps, and a multi-scale attention module, which was used to enhance the feature saliency of objects with different scales, were designed. What's more, to improve the performance of multi-scale object detection, we build 4 different prediction branches for large, medium small and smaller scale objects respectively. At last, we built our own dataset for automatic driving in surface mine called SurMine and test the model at our own datasets and KITTI benchmark. It achieved 82.7 mAP(%) and 92.57 mAP(%) in 32 36 ms on a TITAN RTX, compared to 80.2 mAP(%) and 87.83 mAP(%) in 28 similar to 34 ms by YOLOv7 on SurMine and KITTI benchmarks.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
15  
被引频次合计:  
15  
ISSN:  
2379-8858  
eISSN:  
2379-8904  
入藏号:  
WOS:000981348100025  
Indexed Date:  
2023-06-12  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Management & Control Complex Syst, Beijing 100190, Peoples R China  
Institution Address:  
Waytous Inc, Qingdao 266109, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Artificial Intelligence, Beijing, Peoples R China  
Institution Address:  
North Automatic Control Technol Inst, Taiyuan 030006, Peoples R China  
Institution:  
Institution:  
Waytous Inc  
Institution:  
Institution:  
North Automatic Control Technol Inst

### 586. Research on Multi-Feature Fusion Image Restoration Based on Edge

摘要: Aiming at the lack of rational reasoning of the deep structure inside the image loss area and how to generate more accurate and detailed texture information of images in the current image inpainting field,a multi-feature fusion image inpainting method based on edge conditions is proposed, which named MEGAN(multi-feature fusion network model based on edge condition). The model adopts a two-stage generation idea. Firstly, the edge generation adversarial network is used to repair the edge information of the defective image. Secondly, the complete edge information is used to help the texture detail network to generate a complete image. A gated convolution is added to the generator structure to reduce the interference of invalid pixels on the repair process, and a gated multi-extension convolution block(GM block) is introduced to achieve multi-scale feature extraction of the image to be repaired.The multi-scale spectrally normalized Markov discriminator not only promotes structural coherence and detailed representation of generated images, but also strictly controls the magnitude of gradient variation, thereby improving model accuracy and stabilizing training. The test results on celebA and Places2 datasets show that MEGAN significantly outperforms mainstream image inpainting algorithms in generating reasonable image structure and clear detailed texture.  
摘要:  
针对当前图像修复领域存在的缺乏对图像损失区域深层结构的合理性推理问题,以及如何生成更加准确清晰的纹理信息提出一种基于边缘条件的多特征融合图像修复方法--MEGAN(multi-feature fusion network model based on edge condition)。模型采用两阶段生成思想,使用边缘生成对抗网络修复缺损图像的边缘信息;用完整的边缘信息帮助纹理细节网络生成完整图像。在生成器结构上添加门控卷积以减少无效像素对修复过程的干扰,带门控的多扩张卷积块(gated multi-extension convolution block,GM block)实现对待修复图像的多尺度特征提取。多尺度谱归一化马尔可夫判别器在促进生成图像的结构一致性和细节表现力的同时严格控制梯度变化幅度,从而提高模型精度,稳定训练。在celebA和Places2数据集上的测试结果显示,MEGAN在生成合理的图像结构和准确清晰的细节纹理上明显优于主流的图像修复算法。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7617207  
Indexed Date:  
2024-02-08  
Institution Address:  
湖南工业大学计算机学院, 株洲, 湖南 412000, 中国  
Institution Address:  
School of Computing,Hunan University of Technology, Zhuzhou, Hu'nan 412000  
Institution:  
湖南工业大学计算机学院  
Institution:  
School of Computing,Hunan University of Technology

### 587. Insulator Defect Detection Based on Lightweight Network and Enhanced

摘要: With the development of target detection algorithm embedded in UAV for insulator inspection of transmission towers, a YOLOv5-3S-4PH model based on lightweight network and enhanced multi-scale feature fusion is proposed to detect insulator defects in real time in view of the low detection speed, high network complexity and the difficulty of accurate detection of small defect targets. Firstly, the reconstructed ShuffleNetV2-Stem-SPP(3S) network is used as the backbone of YOLOv5, which reduces the amount of network parameters and calculation significantly. Secondly, the enhanced multi-scale feature fusion network for small targets and four prediction heads(4PH) is added to enhance the networks perception of insulator defects. Combined with Mosaic-9 data enhancement and CIoU loss function, the loss of detection accuracy caused by lightweight is further compensated. Finally, the YOLOv5-3S-4PH model is applied to the self-made insulator dataset for verification. The experimental results show that mean average precision(mAP) is increased by 3%, the detection speed is increased by 81.8%, and parameters and calculation are decreased by 82.4% and 67% compared to original YOLOv5 model. Therefore, the proposed model is more suitable for real-time monitoring of insulator defects deployed on UAV platforms.  
摘要:  
随着无人机搭载目标检测算法在输电杆塔绝缘子巡检领域的发展,针对绝缘子缺陷检测速度较低,网络复杂度高且缺陷小目标难以准确检测的问题,提出一种基于轻量化网络与增强多尺度特征融合的YOLOv5-3S-4PH模型进行绝缘子缺陷实时检测。首先将重构的ShuffleNetV2-Stem-SPP(3S)网络作为YOLOv5的主干网络,显著减小了网络的参数量和计算量;其次引入针对小目标的增强多尺度特征融合网络以及4个预测头,来增强网络对绝缘子缺陷的感知能力,并结合Mosaic-9数据增强、CIoU损失函数进一步补偿轻量化导致的检测精度损失;最后将其应用到自制绝缘子数据集进行验证。实验结果表明,该文所提出的模型相对于未改进的YOLOv5,全类平均精度提高了3%,检测速度提高了81.8%,参数量、计算量分别压缩了82.4%、67%。因此,所提出的模型更适合部署在无人机平台上进行绝缘子缺陷的实时监测。  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
4  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
6  
ISSN:  
1003-6520  
入藏号:  
CSCD:7687293  
Indexed Date:  
2024-06-14  
Institution Address:  
(徐州)中国矿业大学电气工程学院, 徐州, 江苏 221000, 中国  
Institution Address:  
国网江苏省电力有限公司徐州供电分公司, 徐州, 江苏 221000, 中国  
Institution Address:  
School of Electrical Engineering,China University of Mining and Technology(Xuzhou), Xuzhou, Jiangsu 221000, China  
Institution Address:  
Xuzhou Power Supply Company,State Grid Jiangsu Electric Power Co.,Ltd., Xuzhou, Jiangsu 221000, China  
Institution:  
(徐州)中国矿业大学电气工程学院  
Institution:  
国网江苏省电力有限公司徐州供电分公司  
Institution:  
School of Electrical Engineering,China University of Mining and Technology(Xuzhou)  
Institution:  
Xuzhou Power Supply Company,State Grid Jiangsu Electric Power Co.,Ltd.

### 588. LithoSegNet: Regional attention-based deep fusion of multi-scale and

摘要: Lithology identification plays an important role in engineering construction, disaster prediction, reservoir evaluation, and other fields. However, with the development of mechanization and automation, traditional lithology recognition methods are gradually unable to meet existing demands in terms of recognition speed and efficiency. In order to achieve real-time, fast, and accurate identification of lithology, we designed a lightweight model backbone, proposed a multi-scale feature extraction module and a cross -stage feature fusion module based on regional attention, developed a step-wise fusion scheme for multi-stage feature, and developed a real-time lithology segmentation model. The experimental results show that the feature extraction module improves the mIoU of the model by 0.0021; The cross -stage feature fusion module improves the mIoU of the model by 0.0104; The step-wise fusion scheme can effectively fuse shallow spatial features and deep semantic features. The model mIoU using this scheme is 0.0246 and 0.0145 higher than that using SegFormer and SegNeXt fusion schemes, respectively. Our model LithoSegNet can achieve a mIoU of 0.9583 with a speed of 116.25 images per second, comprehensively surpassing many state-of-the-art models. The research results can provide technical support for automated geological sketching in field exploration and on -site construction, and lay the foundation for intelligent construction of underground engineering, which has important scientific and engineering application value.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1365-1609  
eISSN:  
1873-4545  
入藏号:  
WOS:001260331900001  
Indexed Date:  
2024-07-10  
Institution Address:  
Shandong Univ, Geotech & Struct Engn Res Ctr, Jinan 250061, Shandong, Peoples R China  
Institution Address:  
Shandong Univ, Sch Qilu Transportat, Jinan 250061, Shandong, Peoples R China  
Institution Address:  
Shandong Univ, Sch Civil Engn, Jinan 250061, Peoples R China  
Institution:  
Institution:  
Institution:

### 589. An object detection algorithm based on infrared-visible dual modal

摘要: Dual modal image pairs can provide complementary feature information and overcome the limitations of single modal object detection algorithms, thus improving detection performance. In order to fully utilize the features of different modalities, this paper proposes a dual -branch fusion detection network that can simultaneously input both infrared and visible modalities for object detection. The method is based on YOLOv5s, and two attention modules are proposed according to the characteristics of different modalities to enhance the expression capability of infrared and visible image feature information respectively; At the same time, a dual modal fusion module is designed for cross -modal information complementarity by fusing features from corresponding scales of both modalities; Finally, a feature enhancement module is proposed to improve the multi -scale feature fusion capability. The algorithm is validated on the KAIST, FLIR, and GIR datasets and compared with classic single modal and dual modal detection algorithms. Experimental results show that, compared with the baseline algorithm YOLOv5-s alone detecting visible and infrared images separately, the proposed algorithm improves the detection accuracy by 17.6% and 5.9% on the KAIST dataset respectively, and 19.1% and 13% on the FLIR dataset respectively. The proposed algorithm also has a significant advantage in detection accuracy on the self -built GIR dataset.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
ISSN:  
1350-4495  
eISSN:  
1879-0275  
入藏号:  
WOS:001156369200001  
Indexed Date:  
2024-02-12  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Comp Sci & Technol, Xian 710121, Peoples R China  
Institution Address:  
Key Lab Network Data Anal & Intelligent Proc Shaan, Xian 710121, Peoples R China  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Commun & Informat Engn, Xian 710121, Peoples R China  
Institution:  
Institution:  
Key Lab Network Data Anal & Intelligent Proc Shaan  
Institution:

### 590. An adaptive multi-scale feature fusion and adaptive mixture-of-experts

摘要: Health status (HS) assessment and remaining useful life (RUL) prediction are two essential tasks in prognostics and health management. Though HS assessment and RUL prediction based on multi -task learning have achieved considerable progress, the following three challenges still exist. Firstly, it is assumed that these two tasks are related, and the inter -task correlation is ignored. Secondly, it is difficult to extract complex degradation features using the traditional deep network as the shared -bottom structure. Thirdly, inter -task heterogeneity is not considered, which limits the model ' s generalization ability to different tasks. To overcome these challenges, this work proposes an adaptive multi -scale feature fusion and adaptive mixture -of -experts multi -task model for HS assessment and RUL prediction. By exploiting inter -task affinity to measure the correlation between tasks, it provides guidance for multi -task modeling of industrial equipment. Then, an adaptive multi -scale feature fusion mechanism is designed as the shared -bottom structure to improve the feature extraction ability of the multi -task model. Besides, considering inter -task heterogeneity, an adaptive mixture -of -experts mechanism is developed to adaptively fuse expert features for specific tasks, which enhances the model ' s adaptability and generalization to different tasks. Finally, experiments on engine degradation data and tool wear data demonstrate that the proposed model outperforms other state-of-the-art baselines.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0951-8320  
eISSN:  
1879-0836  
入藏号:  
WOS:001240696200001  
Indexed Date:  
2024-06-12  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Civil Aviat, Nanjing 211106, Peoples R China  
Institution:

### 591. Feature Aggregation and Refinement Network for 2D Anatomical Landmark

摘要: Localization of anatomical landmarks is essential for clinical diagnosis, treatment planning, and research. This paper proposes a novel deep network named feature aggregation and refinement network (FARNet) for automatically detecting anatomical landmarks. FARNet employs an encoder-decoder structure architecture. To alleviate the problem of limited training data in the medical domain, we adopt a backbone network pre-trained on natural images as the encoder. The decoder includes a multi-scale feature aggregation module for multi-scale feature fusion and a feature refinement module for high-resolution heatmap regression. Coarse-to-fine supervisions are applied to the two modules to facilitate end-to-end training. We further propose a novel loss function named Exponential Weighted Center loss for accurate heatmap regression, which focuses on the losses from the pixels near landmarks and suppresses the ones from far away. We evaluate FARNet on three publicly available anatomical landmark detection datasets, including cephalometric, hand, and spine radiographs. Our network achieves state-of-the-art performances on all three datasets. Code is available at https://github.com/JuvenileInWind/FARNet.  
在 BIOSIS Citation Index 中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
12  
ISSN:  
0897-1889  
eISSN:  
1618-727X  
入藏号:  
WOS:000885223100001  
PubMed ID:  
36401132  
Indexed Date:  
2022-11-26  
Institution Address:  
Univ Elect Sci & Technol China, Sch Comp Sci & Engn, Chengdu 611731, Sichuan, Peoples R China  
Institution:

### 592. A multiscale feature fusion-guided lightweight semantic segmentation

摘要: Semantic segmentation, a task of assigning class labels to each pixel in an image, has found applications in various real-world scenarios, including autonomous driving and scene understanding. However, its widespread use is hindered by the high computational burden. In this paper, we propose an efficient semantic segmentation method based on Feature Cascade Fusion Network (FCFNet) to address this challenge. FCFNet utilizes a dual-path framework comprising the Spatial Information Path (SIP) and the Context Information Path (CIP). SIP is a shallow structure that captures the local dependencies of each pixel to improve the accuracy of detailed segmentation. CIP is the main branch with a deeper structure that captures sufficient contextual information from input features. Moreover, we design an Efficient Receptive Field Module (ERFM) to enlarge the receptive field in the SIP. Meanwhile, Attention Shuffled Refinement Module is used to refine feature maps from different stages. Finally, we present an Attention-Guided Fusion Module to fuse the low- and high-level feature maps effectively. Experimental results show that our proposed FCFNet achieves 70.7% mean intersection over union (mIoU) on the Cityscapes data set and 68.1% mIoU on the CamVid data set, respectively, with inference speeds of 110 and 100 frames per second (FPS), respectively. Additionally, we evaluated FCFNet on the Nvidia Jetson Xavier embedded device, which demonstrated competitive performance while significantly reducing power consumption.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1556-4959  
eISSN:  
1556-4967  
入藏号:  
WOS:001285601500001  
Indexed Date:  
2024-08-13  
Institution Address:  
Xian Technol Univ, Inst Artificial Intelligence & Data Sci, Xian, Peoples R China  
Institution Address:  
Xian Microelect Technol Inst, Comp Part 3, Xian, Peoples R China  
Institution Address:  
Chaoyue Technol Co Ltd, Tech Management Dept, Xian, Peoples R China  
Institution:  
Institution:  
Xian Microelect Technol Inst  
Institution:  
Chaoyue Technol Co Ltd

### 593. SVSDet: A Fine-Grained Recognition Method for Ship Target Using

摘要: Target recognition from remote sensing images is commonly challenging because of large-scale variations and small objects, and these challenges are more prominent in satellite video images. The current object detection algorithms have some difficulties in fine-grained feature extraction and classification for multiscale and small objects. We propose a novel model called the SVSDet method based on YOLOv5 improvement to address the above-mentioned issues. In this method, we have introduced the space-to-depth module into the backbone of the network, which enhances the network's ability to extract fine-grained features. The neck structure is improved by using the bidirectional feature pyramid network to enhance the network's ability to extract features at multiple scales, thereby improving its overall multiscale feature extraction ability. Subsequently, we have replaced the C3 module in the original network's neck with the C2f module to obtain more abundant gradient flow information. This helps to improve the network's performance further. Finally, the coordinate attention module is introduced into the cross-scale feature connection path, which effectively enhances the network's target detection and recognition performance. We have conducted extensive comparative experiments and ablation experiments on the publicly available datasets ShipRSImageNet and SAT-MTB to confirm the effectiveness of our proposed SVSDet method. The performance of this approach is then evaluated using Jilin 1 satellite video data, and it outperforms the main YOLO series algorithms currently used.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001174290400002  
Indexed Date:  
2024-03-24  
Institution Address:  
China Univ Petr East China, Coll Oceanog & Space Informat, Qingdao 266580, Peoples R China  
Institution:

### 594. Enhancing Measurement Precision for Rotor Vibration Displacement via a

摘要: Recent years have seen the widespread utilization of vision-based noncontact methods for measuring rotor vibrations, but the measurement accuracy of such approaches is still substantially constrained by both the acquisition environment and the equipment, for which improving the quality and clarity of the captured sequence frames would be an effective solution strategy. In this article, a progressive video super-resolution (VSR) reconstruction network is thus constructed to enhance the image feature information during the preliminary phase of vibration displacement measurement, elevating the measurement accuracy while increasing the capture accuracy of the object detection algorithm. To address the challenge of the impractical application of VSR reconstruction methods in diverse industrial conditions, our approach employs pixel displacements between adjacent frames as a reference for motion estimation, ensuring effective feature alignment through a prealignment module. Additionally, a deep feature extraction module is implemented to capture long-range dependencies in multiscale feature representations, crucial for preserving structural image information. To further enhance reconstruction optimization, a feature fusion module (FFM) is introduced, integrating information from diverse rotor images. The experimental results demonstrate that the proposed network surpasses current advanced multiple comparison networks in reconstructing rotor datasets across diverse conditions and rotational speeds and achieves this with a modest parameter count and short run-time, striking a trade-off between computational cost and performance. Specifically, the network proposed in this article achieves peak signal-to-noise ratio (PSNR) values of 41.07, 26.11, 25.05, and 44.96 respectively, with less than half the parameter count of BasicVSR++ across four distinct rotor datasets. In comparison to other VSR networks, the reconstructed image frames in our network exhibit a smooth vibration displacement curve and minimal deviation.  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001200475300012  
Indexed Date:  
2024-05-17  
Institution Address:  
Kunming Univ Sci & Technol, Fac Mech & Elect Engn, Kunming 650500, Peoples R China  
Institution Address:  
Adv Equipment Intelligent Maintenance Engn Res Ctr, Kunming 650500, Peoples R China  
Institution:  
Institution:  
Adv Equipment Intelligent Maintenance Engn Res Ctr

### 595. YOLOv5s-D: A Railway Catenary Dropper State Identification and Small

摘要: High-speed railway catenaries are vital components in railway traction power supply systems. To ensure stable contact between the pantograph and the catenary, droppers are positioned between the messenger wire and contact line. The failure of one or more droppers will affect the power supply of the catenary and the operation of the railway. In this paper, we modify the You Only Look Once version five (YOLOv5) model in several ways and propose a method for improving the identification of dropper status and the detection of small defects. Firstly, to focus on small target features, the selective kernel attention module is added to the backbone. Secondly, the feature graphs of different scales extracted from the backbone network are fed into the bidirectional feature pyramid network for multiscale feature fusion. Thirdly, the YOLO head is replaced by a decoupled head to improve the convergence speed and detection accuracy of the model. The experimental results show that the proposed model achieves a mean average precision of 92.9% on the dropper dataset, an increase of 3.8% over the results using YOLOv5s. The detection accuracy of small dropper defects reaches 79.2%, representing an increase of 10.8% compared with YOLOv5s and demonstrating that our model is better at detecting small defects.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
eISSN:  
2076-3417  
入藏号:  
WOS:001028072400001  
Indexed Date:  
2023-07-28  
Institution Address:  
Beijing Union Univ, Urban Rail Transit & Logist Coll, Beijing 100101, Peoples R China  
Institution Address:  
China Acad Railway Sci Corp Ltd, Inst Comp Technol, Beijing 100081, Peoples R China  
Institution:  
Institution:  
China Acad Railway Sci Corp Ltd

### 596. A low-light image enhancement network based on multi-layer feature

摘要: Current low-light image enhancement methods often suffer from insufficient detail enhancement, color distortion, excessive brightness enhancement, and limited generalization performance. To improve these issues, this paper proposes a multi-layer feature enhancement expression method utilizing multi-branch attention mechanisms. Specifically, single-level or single-scale feature extraction often fails to capture sufficient local details and global contextual features. Therefore, this article designs a multi-scale feature information extension module to obtain rich multi-scale feature information from different levels. Furthermore, to better capture key feature information across different levels, a multi-branch attention module is designed according to the characteristics of features at different scales. For example, lower network layers contain rich detailed information, for which local attention is used to optimize feature extraction. Higher network layers contain a wide range of contextual features that are processed using global attention. By using this strategy to enhance the feature expression ability of images. Finally, these feature information are fused to form a more accurate image feature representation. Extensive experiments on multiple public datasets demonstrate that the proposed method performs well and can enhance low-light images effectively.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001374140100002  
Indexed Date:  
2024-12-16  
Institution Address:  
Putian Univ, Engn Training Ctr, Putian 351100, Peoples R China  
Institution:

### 597. A Highway Traffic Image Enhancement Algorithm Based on Improved GAN in

摘要: To address the problems of low resolution and blurred details in highway images caused by factors such as rain and fog, illumination interference, and nighttime lighting, this paper proposes a highway traffic image enhancement algorithm based on improved GAN in complex weather conditions. The attention mechanism and the multiscale feature fusion were combined to improve the generator network, which could effectively reduce noise while improving the attention of high-frequency region information. The improved PatchGAN in the discriminator used a local discrimination strategy to distinguish the generated image from the real image, and then the Nash equilibrium was achieved through the continuous interaction between the generator and the discriminator, to ensure the integrity and authenticity of the restored image. Compared with other image enhancement algorithms, using PSNR and SSIM as measurement indicators, the experimental results showed that the proposed algorithm's results were, respectively, 21.97% and 12.89% higher in nighttime enhancement, 26.16% and 12.75% higher in rain removal, and 26.56% and 12.1% higher in fog removal. The proposed algorithm can not only retain the image details and feature information, but also produce effective denoising, which increases the reliability of image-based traffic information processing and analysis.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
15  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
16  
ISSN:  
1524-9050  
eISSN:  
1558-0016  
入藏号:  
WOS:000958535700001  
Indexed Date:  
2023-04-14  
Institution Address:  
Changan Univ, Sch Informat Engn, Xian 710064, Peoples R China  
Institution Address:  
Traff Management Res Inst Minist Publ Secur, Wuxi 214151, Peoples R China  
Institution Address:  
Changan Univ, Sch Elect & Control Engn, Xian 710064, Peoples R China  
Institution Address:  
Xian Technol Univ, Sch Elect Informat Engn, Xian 710021, Peoples R China  
Institution:  
Institution:  
Traff Management Res Inst Minist Publ Secur  
Institution:  
Institution:

### 598. A Multilevel Damage Assessment Framework for Mixed-Hazard Buildings With

摘要: After a natural disaster, it is crucial to locate the affected areas and assess building damage for emergency rescue operations. Automatic analysis technology based on remote sensing imagery improves the effectiveness of building damage assessment. However, the existing methods face challenges in determining building damage levels across different types of disasters owing to complex appearance features in remote sensing images. These methods for extracting building damage features are insufficient and lack attention to subtle differences between some visually similar categories. Therefore, this study proposes a multilevel damage assessment (MLDA) universal framework for mixed-hazard buildings. MLDA designs a global spatial feature guidance (GSFG) module to extract high-level semantic features. GSFG pays more attention to the spatial deformation of buildings by long-distance context modeling. It extracts common features of multiple damaged buildings by calculating the deformation similarity under a global perspective, accurately determining their damage levels without considering the impact of the type of disaster. Then, MLDA constructs a damage change feature attention (DCFA) module to enhance heterogeneous features between pre- and postdisaster images. DCFA identifies subtle differences between different damage levels in the process of multiscale feature fusion by channel and spatial dual attention mechanisms. Compared with the state-of-the-art methods, MLDA assesses the damage level of mixed disaster-stricken buildings effectively. Especially for minor damaged buildings, the accuracy rate improved by about 2.3%. The source code is available at https://github.com/DongzheHan-cv/MLDA.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001290229100007  
Indexed Date:  
2024-08-18  
Institution Address:  
Inst Disaster Prevent, Software Engn Technol Res Ctr, Sanhe 065201, Peoples R China  
Institution:

### 599. SiamMaskAttn: inverted residual attention block fusing multi-scale

摘要: Multitask learning combining visual object tracking and other computer vision tasks has received increasing attention from researchers. Among them, the SiamMask algorithm can accomplish both object tracking and object segmentation tasks by utilizing a Siamese backbone network and a three-branch regression head. The mask refinement branch is the core innovation part of the SiamMask, which hierarchically integrates the features of the search region and the tracking correlation score maps. However, SiamMask and its subsequent improved algorithms do not fully integrated the target semantic information contained in multi-scale features into the mask refinement branch. To address the above problems, a module named inverted residual attention block is proposed, which combines the inverted residual structure and channel attention mechanism. The channel attention mechanism can effectively enhance the key information of the object and suppress the background noises by assigning weights to the feature channels output by different convolution kernels, thereby better handling the motion and deformation of the tracking object. Based on the proposed module and spatial attention mechanism, a novel multi-scale feature fusion method of the search region and tracking correlation score maps is proposed. The spatial attention mechanism can help the network focus on the region where the object is located and reduce the sensitivity to background interference, thus improving the accuracy and stability of tracking. Under the condition of using the same hardware and datasets, ablation experiments prove that the proposed improvements for the mask refinement branch are effective. Compared with the baseline SiamMask, the proposed method has achieved comparable segmentation results on the DAVIS datasets with improved speed. The expected average overlap on VOT-2018 has increased by 3.7%. The total number of parameters is reduced by 6.6%, including a 53.2% reduction in the number of parameters in the mask refinement branch.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001099853200001  
Indexed Date:  
2023-12-03  
Institution Address:  
Xihua Univ, Sch Comp & Software Engn, Chengdu 610039, Peoples R China  
Institution:

### 600. GDM-YOLO: A Model for Steel Surface Defect Detection Based on YOLOv8s

摘要: Steel materials are extensively used across various industries. Detecting surface defects in steel strips during production processes is crucial. Existing steel surface defect detection methods exhibit inadequate accuracy and excessive computational complexity, posing challenges for real-time industrial deployment. In this paper, a novel model is designed named GDM-YOLO, specifically tailored for steel surface defect detection tasks, built upon the YOLOv8s network. Firstly, the Space-to-Depth Ghost Convolution (SPDG) downsampling module is introduced and used in the backbone network, aimed at minimizing information loss during downsampling operations while optimizing computational efficiency. Secondly, this work introduces the C2f-Dilated-Reparam-Block (C2f-DRB) module, leveraging reparameterization and large kernel convolutions to enhance feature extraction capabilities without compromising inference costs. Lastly, the novel Multiscale Feature Enhancement Block (MFEB) module was designed, to enhance the small target detection layer by integrating multi-scale feature fusion, further improving detection accuracy. Experimental results demonstrate a 3% improvement in detection accuracy on the NEU-DET dataset compared to the baseline YOLOv8s model. Our approach achieves superior detection performance while reducing parameter requirements and computational complexity, meeting the real-time demands of steel surface defect detection in industrial production.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001337449000001  
Indexed Date:  
2024-10-25  
Institution Address:  
Changchun Univ Technol, Coll Comp Sci & Engn, Changchun 130012, Peoples R China  
Institution Address:  
Changchun Univ Technol, Grad Sch Changchun Univ Technol, Changchun 130012, Peoples R China  
Institution:  
Institution:

### 601. Object Recognition for Millimeter Wave MIMO-SAR Images Based on

摘要: There are several complex situations in recognizing concealed objects from millimeter wave multiple-input multiple-output synthetic aperture radar (MIMO-SAR) security images, such as incomplete imaging of objects, partially occluded objects, and overlapping objects, which are detrimental to the accurate recognition of concealed objects. To solve these problems, a concealed object detection method based on a high-resolution feature recursive alignment fusion network (HR-FRAFnet) is proposed. The HR-FRAFnet can segment the object area from the grayscale image with complex human background and complete the recognition. The overall architecture of the HR-FRAFnet follows the encoder-decoder framework. Specifically, in the encoder stage, a deep parallel feature extraction network (DPFEN) connects the multiresolution feature maps in parallel and repeats multiscale feature fusion. This approach suppresses the background noise flowing and retains more recognizable target characteristics. Then, in the decoder stage, a feature recursive alignment fusion module (FRAFM) is designed to enhance the perception of object edges. The FRAFM effectively improves the segmentation accuracy of objects while decreasing the computational complexity of the network. Besides, we employ a combined loss function to alleviate the foreground-background imbalance problem in MIMO-SAR images. Homemade human security screening image datasets are used for evaluation. The experimental results show that the proposed method outperforms existing semantic segmentation methods in mean intersection over union (mIoU) and reduces the incidence of missed and error detection of targets.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
被引频次合计:  
9  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001030784400116  
Indexed Date:  
2023-08-16  
Institution Address:  
Univ Shanghai Sci & Technol, Terahertz Technol Innovat Res Inst, Shanghai 200093, Peoples R China  
Institution:

### 602. RT-YOLO: A Residual Feature Fusion Triple Attention Network for Aerial

摘要: In recent years, target detection of aerial images of unmanned aerial vehicle (UAV) has become one of the hottest topics. However, target detection of UAV aerial images often presents false detection and missed detection. We proposed a modified you only look once (YOLO) model to improve the problems arising in object detection in UAV aerial images: (1) A new residual structure is designed to improve the ability to extract features by enhancing the fusion of the inner features of the single layer. At the same time, triplet attention module is added to strengthen the connection between space and channel and better retain important feature information. (2) The feature information is enriched by improving the multi-scale feature pyramid structure and strengthening the feature fusion at different scales. (3) A new loss function is created and the diagonal penalty term of the anchor frame is introduced to improve the speed of training and the accuracy of reasoning. The proposed model is called residual feature fusion triple attention YOLO (RT-YOLO). Experiments showed that the mean average precision (mAP) of RT-YOLO is increased from 57.2% to 60.8% on the vehicle detection in aerial image (VEDAI) dataset, and the mAP is also increased by 1.7% on the remote sensing object detection (RSOD) dataset. The results show that the RT-YOLO outperforms other mainstream models in UAV aerial image object detection.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001126979300003  
Indexed Date:  
2024-01-13  
Institution Address:  
Hengyang Normal Univ, Coll Comp Sci & Technol, Hengyang 421002, Peoples R China  
Institution:

### 603. Research on Multi-Scale Feature Fusion Dehazing Network Based on Feature

摘要: Haze,formed by the accumulation and concentration of atmospheric pollutants under meteorological conditions,such as temperature inversion,severely limits visibility.Image dehazing techniques aim to eliminate issues caused by haze,such as image blur and low contrast,thereby enhancing image clarity and visibility.However,challenges persist regarding the loss of image details.To address this issue,a feature difference-based multi-scale feature fusion dehazing network known as FD-CA dehaze is proposed in this study.In this network,the basic block structure of the FFA-Net is enhanced by extracting intermediate feature information from the feature difference,coordinate,and channel dimensions.An Effective Coordinate Attention (ECA) module that combines global pooling,max pooling,and coordinate positional information is introduced.This module mitigates the positional information loss during feature fusion.By integrating channel attention with the ECA module,a Dual Attention (D-CA) model that enables better utilization of spatial and channel information is constructed.Consequently,the model exhibits enhanced performance in image dehazing tasks.Furthermore,the loss function is improved by combining L1 loss function with perceptual loss.Experimental results on the Synthetical Objective Test Set (SOTS) and Hybrid Subjective Test Set (HSTS) demonstrate that the FD-CA dehaze network achieves a Peak Signal-to-Noise Ratio (PSNR) of 37.93 dB and a Structural Similarity Index (SSIM) of 0.990 5.Experimental results demonstrate that compared to classic dehazing networks such as FFA-Net and GridDehazeNet,FD-CA dehaze achieves significant improvement and better dehazing performance.  
摘要:  
雾霾是大气污染物在逆温等气象条件下聚集和累积而形成的混浊物质,其可见性十分有限。图像去雾技术能够消除由雾霾导致的模糊、低对比度等问题,提高图像的清晰度和可见性,但也存在图像细节信息丢失等问题。为此,提出一种基于特征差异的多尺度特征融合去雾(FD-CA dehaze)网络。对FFA-Net的基本块结构进行改进,分别从特征差异维度、坐标维度和通道维度提取中间特征信息。提出有效坐标注意力(ECA)模块,将全局池化、最大池化与坐标位置信息相结合,用于减轻特征融合过程中的位置信息丢失问题;将通道注意力与ECA模块相结合,构建双注意力(D-CA)模型,更好地利用空间信息和通道信息,进一步提升模型在图像去雾任务中的表现。在此基础上,改进损失函数,将L1损失与感知损失相结合。在综合目标测试集(SOTS)和混合主观测试集(HSTS)中进行实验,结果表明,FD-CA dehaze网络在峰值信噪比、结构相似度2个指标上分别达到37.93 dB和0.990 5,相较于FFA-Net、GridDehazeNet等经典去雾网络,FD-CA dehaze的去雾效果得到明显提升。  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1000-3428  
入藏号:  
CSCD:7698348  
Indexed Date:  
2024-06-29  
Institution Address:  
中北大学软件学院, 太原, 山西 030051, 中国  
Institution Address:  
School of Software,North University of China, Taiyuan, Shanxi 030051, China  
Institution:  
中北大学软件学院  
Institution:  
School of Software,North University of China

### 604. Urtnet: an unstructured feature fusion network for real-time detection

摘要: Minimally invasive surgery (MIS) is increasingly popular due to its smaller incisions, less pain, and faster recovery. Despite its advantages, challenges like limited visibility and reduced tactile feedback can lead to instrument and organ damage, highlighting the need for precise instrument detection and identification. Current methods face difficulties in detecting multi-scale targets and are often disrupted by blurring, occlusion, and varying lighting conditions during surgeries. Addressing these challenges, this paper introduces URTNet, a novel unstructured feature fusion network designed for the real-time detection of multi-scale surgical instruments in complex environments. Initially, the paper proposes a Stair Aggregation Network (SAN) to efficiently merge multi-scale information, minimizing detail loss in feature fusion and improving detection of blurred and obscured targets. Subsequently, a Multi-scale Feature Weighted Fusion (MFWF) approach is presented to tackle significant scale variations in detection objects and reconstruct the detection layers based on target sizes within endoscopic views. The effectiveness of URTNet is validated through tests on the public laparoscopic dataset m2cai16-tool and another dataset from Sun Yat-sen University Cancer Center, where URTNet achieved average precision scores (AP0.5\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$AP\_{0.5}$$\end{document}) of 93.3% and 97.9%, surpassing other advanced methodologies.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001348452200001  
Indexed Date:  
2024-11-12  
Institution Address:  
Guangdong Univ Technol, Sch Automat, Guangzhou 510006, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Sch Intelligent Syst Engn, Shenzhen 518107, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Dept Neurosurg, Canc Ctr, Guangzhou 510006, Peoples R China  
Institution Address:  
Zhejiang Chinese Med Univ, Zhejiang Prov Hosp Chinese Med, Affiliated Hosp 1, Dept Gastroenterol, Hangzhou 310006, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 605. Heterogeneous Noise Iris Segmentation Based on Attention Mechanism and

摘要: Current segmentation methods do not account for the segmentation accuracy and efficiency of visible and near- infrared heterogeneous noisy iris images; thus, in this study, we propose a codec network based on the fusion of attention mechanism and dense multiscale features. First, an improved residual bottleneck element based on deep separable convolution was introduced to reduce the number of parameters and computation while preventing information loss and gradient confusion. Second, the dense void space pyramid module's void rate combination was improved and placed behind the encoder to improve multiscale feature fusion. Finally, to improve the resolution of noise targets and iris pixels, an efficient parallel space-channel attention module was designed and integrated into each down sampling layer and decoder. The experiments conducted on three open iris data sets show that both the average F1-score and mean intersection over union (mIoU) of the proposed network are superior to the existing algorithms. Compared with the benchmark network, the occupied space, number of parameters, and amount of computation are reduced by 41%, 41. 77%, and 65. 35%, respectively. It can effectively improve the segmentation performance of the network for multispectral noise iris and is easier to be deployed on mobile devices, which can more efficiently and accurately distinguish noise and iris targets.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
ISSN:  
1006-4125  
入藏号:  
WOS:000823174500011  
Indexed Date:  
2022-07-24  
Institution Address:  
Rocket Force Univ Engn, Coll Operat Support, Xian 710025, Shaanxi, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Coll Microelect, Xian 710049, Shaanxi, Peoples R China  
Institution:  
Institution:

### 606. Underwater object detection algorithm based on channel attention and

摘要: Due to the color deviation, low contrast and fuzzy object in underwater optical images, there are some problems in underwater object detection, such as missed detection and false detection. In order to solve the abovementioned problems, an underwater object detection algorithm is proposed based on the channel attention and feature fusion for underwater optical images. The excitation residual module is designed based on the channel attention, and the forward propagation feature information is adaptively allocated weights to highlight the salience of different channel feature maps, which improves the network ability to extract high-frequency information from the underwater images. The multi-scale feature fusion module is designed to add a large scale feature map for object detection, which improves the detection performance of the network for small size objects by using its corresponding small size receptive field, and further improves the detection accuracy of the network for different size objects in the underwater environment. To improve the generalization performance of the network to the underwater environment, the data augmentation method based on the stitching and fusion is designed to simulate the overlap, occlusion and blurring of underwater objects, which improves the adaptability of the network to the underwater environment. Through experiments on the public dataset URPC, the algorithm in this paper improves the mean average precision by 5.42%, 3.20% and 0.9% compared with YOLOv3, YOLOv4 and YOLOv5, respectively, effectively improving the missed and false detection of objects of different sizes in complex underwater environments.  
摘要:  
水下光学图像存在色偏、低对比度、目标模糊的现象,导致水下目标检测时存在漏检、误检等问题。针对上述问题,提出了一种基于通道注意力与特征融合的水下目标检测算法。基于通道注意力设计了激励残差模块,将前向传播的特征信息进行自适应分配权重,以突出不同通道特征图的显著性,提高了网络对水下图像高频信息的提取能力;设计了多尺度特征融合模块,增加了大尺度特征图用于目标检测,利用其对应的小尺度感受野提高了网络对小尺寸目标的检测性能,进一步提高了网络对水下不同尺寸目标的检测精度;为提高网络对水下环境的泛化性能,设计了基于拼接和融合的数据增强方法,模拟水下目标的重叠、遮挡和模糊情况,增强了网络对水下环境的适应性。通过在公共数据集URPC上的实验,与YOLOv3、YOLOv4和YOLOv5相比,所提算法的平均精度均值分别提升5.42%,3.20%和0.9%,有效改善了水下复杂环境中不同尺寸目标漏检、误检的问题。  
在中国科学引文数据库中的被引频次:  
5  
Web of Science 核心合集中的 "被引频次":  
5  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
ISSN:  
1000-2758  
入藏号:  
CSCD:7200764  
Indexed Date:  
2022-07-08  
Institution Address:  
天津城建大学计算机与信息工程学院, 天津 300384, 中国  
Institution Address:  
School of Computer and Information Engineering, Tianjin Chengjian University, Tianjin 300384, China  
Institution:  
天津城建大学计算机与信息工程学院  
Institution:  
School of Computer and Information Engineering, Tianjin Chengjian University

### 607. Accurate UAV Small Object Detection Based on HRFPN and EfficentVMamba

摘要: (1) Background: Small objects in Unmanned Aerial Vehicle (UAV) images are often scattered throughout various regions of the image, such as the corners, and may be blocked by larger objects, as well as susceptible to image noise. Moreover, due to their small size, these objects occupy a limited area in the image, resulting in a scarcity of effective features for detection. (2) Methods: To address the detection of small objects in UAV imagery, we introduce a novel algorithm called High-Resolution Feature Pyramid Network Mamba-Based YOLO (HRMamba-YOLO). This algorithm leverages the strengths of a High-Resolution Network (HRNet), EfficientVMamba, and YOLOv8, integrating a Double Spatial Pyramid Pooling (Double SPP) module, an Efficient Mamba Module (EMM), and a Fusion Mamba Module (FMM) to enhance feature extraction and capture contextual information. Additionally, a new Multi-Scale Feature Fusion Network, High-Resolution Feature Pyramid Network (HRFPN), and FMM improved feature interactions and enhanced the performance of small object detection. (3) Results: For the VisDroneDET dataset, the proposed algorithm achieved a 4.4% higher Mean Average Precision (mAP) compared to YOLOv8-m. The experimental results showed that HRMamba achieved a mAP of 37.1%, surpassing YOLOv8-m by 3.8% (Dota1.5 dataset). For the UCAS\_AOD dataset and the DIOR dataset, our model had a mAP 1.5% and 0.3% higher than the YOLOv8-m model, respectively. To be fair, all the models were trained without a pre-trained model. (4) Conclusions: This study not only highlights the exceptional performance and efficiency of HRMamba-YOLO in small object detection tasks but also provides innovative solutions and valuable insights for future research.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001287106700001  
PubMed ID:  
39124013  
Indexed Date:  
2024-08-15  
Institution Address:  
Wuhan Business Univ, Sch Informat Engn, Wuhan 430056, Peoples R China  
Institution Address:  
Tianjin Univ Technol, Key Lab Comp Vis & Syst, Minist Educ, Tianjin 300384, Peoples R China  
Institution Address:  
Wuhan Coll, Sch Informat Engn, Wuhan 430212, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Elect Informat, Wuhan 430072, Peoples R China  
Institution Address:  
Tianjin Univ Technol, Sch Comp Sci & Engn, Tianjin 300384, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 608. Real-time pixel-wise grasp affordance prediction based on multi-scale

摘要: Purpose This paper aims to use fully convolutional network (FCN) to predict pixel-wise antipodal grasp affordances for unknown objects and improve the grasp detection performance through multi-scale feature fusion. Design/methodology/approach A modified FCN network is used as the backbone to extract pixel-wise features from the input image, which are further fused with multi-scale context information gathered by a three-level pyramid pooling module to make more robust predictions. Based on the proposed unify feature embedding framework, two head networks are designed to implement different grasp rotation prediction strategies (regression and classification), and their performances are evaluated and compared with a defined point metric. The regression network is further extended to predict the grasp rectangles for comparisons with previous methods and real-world robotic grasping of unknown objects. Findings The ablation study of the pyramid pooling module shows that the multi-scale information fusion significantly improves the model performance. The regression approach outperforms the classification approach based on same feature embedding framework on two data sets. The regression network achieves a state-of-the-art accuracy (up to 98.9%) and speed (4 ms per image) and high success rate (97% for household objects, 94.4% for adversarial objects and 95.3% for objects in clutter) in the unknown object grasping experiment. Originality/value A novel pixel-wise grasp affordance prediction network based on multi-scale feature fusion is proposed to improve the grasp detection performance. Two prediction approaches are formulated and compared based on the proposed framework. The proposed method achieves excellent performances on three benchmark data sets and real-world robotic grasping experiment.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
ISSN:  
0143-991X  
eISSN:  
1758-5791  
入藏号:  
WOS:000736637500001  
Indexed Date:  
2022-01-07  
Institution Address:  
Harbin Inst Technol, State Key Lab Robot & Syst, Harbin, Peoples R China  
Institution:

### 609. A multi-layer mask recognition method for Tangut characters

摘要: Aiming at the problem of poor recognition ability of existing methods for fuzzy and mutilated Tangut characters, a Tangut character recognition model MMSFTR is proposed. Firstly, a multilayer mask learning strategy is introduced to extract key character features in a hierarchical manner, assisting the model in understanding the internal structure of the Tangut characters more efficiently, and improving its ability to describe complex features of Tangut characters. Secondly, a multi-scale feature fusion module is designed to extract richer multi-scale features. Then, a channel adaptive attention module is proposed to better select and focus on information from specific channels. A mask attention module is also designed to improve the model's perception capabilities. Finally, a feature enhancement module is designed to optimize multi-level features of the network and enhance deep-level features. Through the collaborative work of these 4 modules, MMSFTR achieves the desired results. Experimental results show that MMSFTR achieves a recognition accuracy of 99.40% on the TCD-E dataset, effectively enhancing the recognition effect of fuzzy and mutilated Tangut characters.  
摘要:  
针对现有方法对模糊、残缺西夏文字识别能力较差的问题,提出西夏文字识别模型MMSFTR 。首先,提出多层掩码学习策略,分层次提取字符关键特征,帮助模型更有效地理解西夏文字内部结构,提高对复杂西夏文字的特征描述能力。其次,设计多尺度特征融合模块,以提取更丰富的多尺度特征。然后,提出通道自适应注意力模块,更好地选择和关注特定通道的信息,并设计掩码注意力模块改善模型感知能力。最后,设计特征增强模块,对网络进行多层次特征优化,并进行深层次特征增强。MMSFTR通过4个模块的协同作业,使得模型达到了预期效果。实验结果显示:MMSFTR在TCD-E数据集上达到99.40%的识别准确率,有效提升了对模糊、残缺西夏文字的识别效果。  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1007-130X  
入藏号:  
CSCD:7868164  
Indexed Date:  
2025-03-01  
Institution Address:  
北方民族大学计算机科学与工程学院;;图像图形智能信息处理国家民委重点实验室, ;;图像图形智能信息处理国家民委重点实验室, 银川;;银川, 宁夏;;宁夏 750021;;750021, 中国  
Institution Address:  
北方民族大学计算机科学与工程学院, 银川, 宁夏 750021, 中国  
Institution Address:  
北方民族大学数学与信息科学学院, 银川, 宁夏 750021, 中国  
Institution Address:  
College of Computer Science and Engineering, North Minzu University;;The Key Laboratory of Images and Graphics Intelligent Processing of State Ethnic Affairs Commission, ;;The Key Laboratory of Images and Graphics Intelligent Processing of State Ethnic Affairs Commission, Yinchuan;;Yinchuan, ;; 750021;;750021  
Institution Address:  
College of Computer Science and Engineering, North Minzu University, Yinchuan, Ningxia 750021, China  
Institution Address:  
School of Mathematics and Information Science, North Minzu University, Yinchuan, Ningxia 750021, China  
Institution:  
北方民族大学计算机科学与工程学院;;图像图形智能信息处理国家民委重点实验室  
Institution:  
北方民族大学计算机科学与工程学院  
Institution:  
北方民族大学数学与信息科学学院  
Institution:  
College of Computer Science and Engineering, North Minzu University;;The Key Laboratory of Images and Graphics Intelligent Processing of State Ethnic Affairs Commission  
Institution:  
College of Computer Science and Engineering, North Minzu University  
Institution:  
School of Mathematics and Information Science, North Minzu University

### 610. Feature fusion for object detection at one map

摘要: The scale feature plays a crucial role in the detector, and existing methods adopt the feature pyramid based on multiple maps. This paper focuses on a single map and proposes an encoder called SFMF which can employ multi-scale feature fusion on a map. One of the crucial techniques underlying SFMF is a fine-grained weighting method that is used to fast discard unneeded pixel channels during the fusion process. YOLOF (you only look one-level feature) with SFMF (single feature map fusion) achieve 38.5 mAP in the ResNet50 and 40.3 mAP in the ResNet101, which improves 0.8 and 0.5 mAP than the baseline, respectively. Meta-ACON is used to auto learn activate the neurons or not in the backbone. With the Meta-ACON and SFMF, YOLOF can achieve 39.1 and 40.4 mAP, surpassing the baseline by 1.4 and 0.6 mAP on COCO val-dev. In addition, YOLOF with SFMF achieves 54.8 mAP, improving the performance by an absolute 4.9 mAP on the aircraft detection dataset, with a slight sacrificing efficiency (1 FPS) in inference. (c) 2022 Published by Elsevier B.V.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
8  
ISSN:  
0262-8856  
eISSN:  
1872-8138  
入藏号:  
WOS:000799606300002  
Indexed Date:  
2022-06-03  
Institution Address:  
Guangdong Univ Technol, Guangzhou 510006, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Shanghai, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, NingBo Artificial Intelligent Inst, Shanghai, Peoples R China  
Institution:  
Institution:  
Institution:

### 611. Methane concentration inversion under multiple conditions using feature

摘要: In the realm of methane concentration measurement using tunable diode laser absorption spectroscopy (TDLAS), relying on a single second harmonic signal peak often leads to information loss. Furthermore, second harmonic peaks exhibit conflicts when subjected to different parameters, thereby posing challenges for accurate methane concentration measurement. This paper presents a novel approach employing a multi -scale feature fusion residual network (MResNet) to measure methane concentration under various operational conditions. To begin, we established a hardware experimental platform based on TDLAS to collect direct absorption spectrum data, and second harmonic data with varying parameters. Then we conducted a comparative analysis between MResNet and conventional models. The results of our experiments demonstrate the effectiveness of MResNet in predicting methane concentration, with a decrease in MAE by 39% and MSE by 70%. Notably, the coefficient of determination for methane concentration prediction using MResNet both reached 99.92% when applied to both direct absorption spectrum and second harmonic data. This capability empowers accurate methane concentration measurement across a wide range of operating conditions.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0030-4018  
eISSN:  
1873-0310  
入藏号:  
WOS:001223982000001  
Indexed Date:  
2024-05-24  
Institution Address:  
Northeast Petr Univ, Coll Elect & Informat Engn, Daqing 163318, Peoples R China  
Institution:

### 612. Belt conveyor idler fault diagnosis method based on multi-scale feature

摘要: Existing idler fault diagnosis methods have problems in failing to fully obtain global context information and providing poor diagnostic accuracy. To address these problems, this paper investigates a new method for diagnosing faults in belt conveyor idlers, based on analysis of their acoustic signals. The method is also applied to existing databases of bearing fault data. Firstly, an eight -element microphone array sound signal collector is designed to suppress environmental noise and raise the signal-to-noise ratio of the idler sound signal. Secondly, a multi -scale feature fusion (MSFF) module is constructed to learn complementary information between features at different scales. Then, a residual mask convolutional attention (MCA) module is designed to raise the modelling capability of local features and global contextual information. Finally, the structure of the ResNet-18 network is optimised to improve model fitting performance. Experimental results on self-made and public datasets show that the suggested method outperforms other comparative methods, real-time accurate detection and classification of belt idler faults and faults.  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1354-2575  
eISSN:  
1754-4904  
入藏号:  
WOS:001240080600005  
Indexed Date:  
2024-06-15  
Institution Address:  
Tiangong Univ, Sch Elect & Informat Engn, Tianjin 300387, Peoples R China  
Institution Address:  
Tianjin Key Lab Optoelect Detect Technol & Syst, Tianjin 300387, Peoples R China  
Institution Address:  
Tiangong Univ, Ctr Engn Internship & Training, Tianjin 300387, Peoples R China  
Institution:  
Institution:  
Tianjin Key Lab Optoelect Detect Technol & Syst  
Institution:

### 613. DHIQA: Quality assessment of dehazed images based on attentive

摘要: Haze is a ubiquitous atmospheric phenomenon that seriously influences the visibility of images. To this end, numerous image dehazing models have been proposed to improve the visual quality of hazy images. However, the quality assessment of dehazed images has fallen behind. Traditional IQA methods for dehazed images rely on complex handcrafted features that may not accurately evaluate the quality of dehazed images. In this paper, we present a novel no-reference image quality assessment network for dehazed images (called DHIQA). Firstly, we propose a multi-scale feature fusion network that considers typical dehazed distortion related features, such as contrast, dark channel, edge, etc. Then, inspired by the contrast sensitivity function of the Human Visual System (HVS), we propose a contrast attentive module to enhance the effectiveness of the proposed IQA network. Moreover, in addition to the absolute quality score of dehazed images, their rankings are also important for quality assessment. Thus, we propose a new ranking loss, which takes both the absolute quality score and their rankings into consideration. Extensive experimental results on six publicly available dehazed IQA databases show that the proposed model achieves superior performance compared to current state-of-the-art metrics.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
0141-9382  
eISSN:  
1872-7387  
入藏号:  
WOS:001059099100001  
Indexed Date:  
2023-09-19  
Institution Address:  
Shenzhen Univ, Shenzhen Key Lab Adv Machine Learning & Applicat, Shenzhen 518060, Peoples R China  
Institution Address:  
Shenzhen Univ, Guangdong Key Lab Intelligent Informat Proc, Shenzhen 518060, Peoples R China  
Institution Address:  
Shenzhen Univ, Inst Artificial Intelligence & Adv Commun, Shenzhen 518060, Peoples R China  
Institution Address:  
Shenzhen Univ, Coll Elect & Informat Engn, Shenzhen 518060, Peoples R China  
Institution Address:  
INSA Rennes, Natl Inst Appl Sci Rennes, Rennes, France  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 614. Multiscale ship detection based on cascaded dense weighted networks in

摘要: The sizes of ships vary by type in synthetic aperture radar (SAR) images, and this multiscale problem degrades the accuracy of ship detection. We present a multiscale ship detection method based on cascaded dense weighted networks in SAR images. Dense weighted fusion is performed in the feature extraction module and feature fusion module by cascading, i.e., feature weighting is achieved by embedding a convolutional block attention module after feature layers at different scales, followed by dense connection to fuse features at different scales. This makes it possible to refine the cascaded feature mapping, enhance the information transfer of shallow and deep features, and build a multiscale feature representation model. Experiments on the extended high-resolution SAR images dataset show that the accuracy of this method is 74.87% in inshore scenes, and 97.39% in offshore scenes, which exceeds that of other ship detection methods. In addition, in two SAR images from TerraSAR-X, the detection accuracy of multiscale ships reaches 86.63%, which indicates the method's better robustness. (c) 2022 Society of Photo-Optical Instrumentation Engineers (SPIE)  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
1931-3195  
入藏号:  
WOS:000892785000031  
Indexed Date:  
2022-12-19  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Astronaut, Nanjing, Peoples R China  
Institution Address:  
Nat Key Lab Sci & Technol Space Microwave, Xian, Peoples R China  
Institution:  
Institution:  
Nat Key Lab Sci & Technol Space Microwave

### 615. Pedestrian re-identification method based on multi-scale feature fusion

摘要: Pedestrian re-identification tasks lack the consideration of the pedestrian feature scale variation during feature extraction,so that they are easily affected by environment and have low accuracy of pedestrian re-identification. In order to solve the problem,a pedestrian re-identification method based on multi-scale feature fusion was proposed. Firstly, in the shallow layer of the network,multi-scale pedestrian features were extracted through mixed pooling operation,which was helpful to improve the feature extraction capability of the network. Then,strip pooling operation was added to the residual block to extract the remote context information in horizontal and vertical directions respectively,which avoided the interference of irrelevant regions. Finally,after the residual network,the dilated convolutions with different scales were used to further preserve the multi-scale features,so as to help the model to analyze the scene structure flexibly and effectively. Experimental results show that,on Market-1501 dataset,the proposed method has the Rank1 of 95.9%,and the mean Average Precision(mAP)of 88.5%;on DukeMTMC-reID dataset,the proposed method has the Rank1 of 90.1%,and the mAP of 80.3%. It can be seen that the proposed method can retain the pedestrian feature information better,thereby improving the accuracy of pedestrian re-identification tasks.  
摘要:  
针对行人重识别任务在特征提取时缺乏对行人特征尺度变化的考虑,导致其易受环境影响而具有低行人重识别准确率的问题,提出了一种基于多尺度特征融合的行人重识别方法。首先,在网络浅层通过混合池化操作来提取多尺度的行人特征,从而帮助网络提升特征提取能力;然后,在残差块内添加条形池化操作以分别提取水平和竖直方向的远程上下文信息,从而避免无关区域的干扰;最后,在残差网络之后利用不同尺度的空洞卷积进一步保留多尺度的特征,从而帮助模型灵活有效地解析场景结构。实验结果表明,在Market-1501数据集上,所提方法的Rank1达到95.9%,平均精度均值(mAP)为88.5%;在DukeMTMC-reID数据集上,该方法的Rank1达到90.1%,mAP为80.3%。可见所提方法能够较好地保留行人特征信息,从而提高行人重识别任务准确率。  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
被引频次合计:  
3  
ISSN:  
1001-9081  
入藏号:  
CSCD:7073394  
Indexed Date:  
2021-12-31  
Institution Address:  
山西大学计算机与信息技术学院, 太原, 山西 030006, 中国  
Institution Address:  
School of Computer and Information Technology, Shanxi University, Taiyuan, Shanxi 030006, China  
Institution:  
山西大学计算机与信息技术学院  
Institution:  
School of Computer and Information Technology, Shanxi University

### 616. Unsupervised Segmentation Algorithm Based on Multi-Scale Feature Fusion

摘要: In the process of intelligent upgrading of factories, there are many application scenarios that need to use semantic segmentation algorithms. However, the use of fully supervised semantic segmentation methods requires a lot of labor cost for sample labeling, so it is necessary to study unsupervised semantic segmentation methods. Aiming at the semantic segmentation of characters imprinted by graphite electrodes in a local carbon factory, an unsupervised semantic segmentation method CycleGAN-Seg is proposed. Combining the idea of cross-layer connection and atrous spatial pooling pyramid(ASPP), a novel multi-scale feature fusion generator is constructed, and an improved attention module is added to improve the network performance. At the same time, a new U-shaped discriminator is proposed to discriminate the reconstructed images. In the semantic segmentation experiment of the imprinted character dataset on the graphite electrode surface, the MIoU value can reach 70.81%. The semantic segmentation effect basically meets the recognition needs, and it is expected to replace the fully supervised learning method in this industrial scenario to save the cost of manual annotation and achieve the purpose of rapid training and deployment.  
摘要:  
工厂在智能化升级过程中,有很多应用场景需要用到语义分割。然而使用全监督语义分割方法需要耗费大量人力成本进行样本标注,所以研究无监督语义分割方法很有必要。针对本地某碳素厂石墨电极压印字符的语义分割问题,提出了一种无监督语义分割方法CycleGAN-Seg。结合跨层连接和空洞空间池化金字塔(ASPP)的思想,构建了新型多尺度特征融合生成器,加入了改进的注意力模块以提升网络性能。同时提出一种新的U形判别器对重构图像进行判别。在石墨电极表面压印字符数据集语义分割实验中,MIoU值可达70.81%,分割效果基本满足识别需要,有望在该工业场景中替代全监督学习方法,以节省人工标注成本,达到快速训练和部署的目的。  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7450737  
Indexed Date:  
2023-07-28  
Institution Address:  
河南大学物理与电子学院, 开封, 河南 475004, 中国  
Institution Address:  
开封平煤新型炭材料科技有限公司, 开封, 河南 475004, 中国  
Institution Address:  
School of Physics and Electronics,Henan University, Kaifeng, He'nan 475004  
Institution Address:  
Kaifeng Pingmei New Carbon Materials Technology Co.,Ltd., Kaifeng, He'nan 475004  
Institution:  
河南大学物理与电子学院  
Institution:  
开封平煤新型炭材料科技有限公司  
Institution:  
School of Physics and Electronics,Henan University  
Institution:  
Kaifeng Pingmei New Carbon Materials Technology Co.,Ltd.

### 617. ESL-YOLO: Small Object Detection with Effective Feature Enhancement and

摘要: Improving the detection of small objects in remote sensing is essential for its extensive use in various applications. The diminutive size of these objects, coupled with the complex backgrounds in remote sensing images, complicates the detection process. Moreover, operations like downsampling during feature extraction can cause a significant loss of spatial information for small objects, adversely affecting detection accuracy. To tackle these issues, we propose ESL-YOLO, which incorporates feature enhancement, fusion, and a local attention pyramid. This model includes: (1) an innovative plug-and-play feature enhancement module that incorporates multi-scale local contextual information to bolster detection performance for small objects; (2) a spatial-context-guided multi-scale feature fusion framework that enables effective integration of shallow features, thereby minimizing spatial information loss; and (3) a local attention pyramid module aimed at mitigating background noise while highlighting small object characteristics. Evaluations on the publicly accessible remote sensing datasets AI-TOD and DOTAv1.5 indicate that ESL-YOLO significantly surpasses other contemporary object detection frameworks. In particular, ESL-YOLO enhances mean average precision mAP by 10% and 1.1% on the AI-TOD and DOTAv1.5 datasets, respectively, compared to YOLOv8s. This model is particularly adept at small object detection in remote sensing imagery and holds significant potential for practical applications.  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:001377594800001  
Indexed Date:  
2024-12-21  
Institution Address:  
Natl Lab Adapt Opt, Chengdu 610209, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 101408, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Opt & Elect, Chengdu 610209, Peoples R China  
Institution:  
Natl Lab Adapt Opt  
Institution:  
Institution:

### 618. Research on Object Detection Algorithm for Remote Sensing Images Based

摘要: In order to tackle the challenges of complex background, feature conflict and variable target scale in remote sensing images, a target detection method based on multi-scale feature fusion using YOLOv8s as the baseline model is proposed for remote sensing images. Firstly, RepVGG network is used as the feature extraction network to improve the feature extraction capability and effectively achieve the capture of global semantic information. Secondly, a tripolar integrative fusion (TIF) module is designed in the neck network to improve the detection accuracy of targets at all scales through the effective fusion of positional and semantic information. Finally, the SlideLoss function is used as the classification loss function to enhance the detection ability of difficult targets and enhance the precision of target detection. The experimental results indicate that the improved model achieves 94.4%, 93.0% and 95.5% detection accuracies on the NWPU VHR-10, RSOD and UCAS-AOD datasets respectively, which are 5.1, 6.0 and 4.4?percentape points higher than the baseline model, which is superior to the other methods in terms of accuracy and is able to better complete target detection task in remote sensing images.  
摘要:  
针对遥感图像背景复杂、特征冲突以及目标尺度多变等问题,以YOLOv8s为基线模型,提出一种基于多尺度特征融合的遥感图像目标检测方法。将RepVGG网络作为模型的特征提取网络,提高模型的特征提取能力,有效实现全局语义信息的捕获;在颈部网络提出了一种三极综合性融合模块(tripolar integrative fusion,TIF),通过位置信息和语义信息的有效融合,提高对各个尺度目标的检测精度;将SlideLoss滑动损失函数作为模型的分类损失函数,增强模型对困难目标的检测能力,从而提高目标检测的准确性。实验结果表明,改进后的模型在NWPU VHR-10、RSOD和UCAS-AOD数据集上分别取得了94.4%、93.0%和95.5%的检测精度,相较基线模型分别提升了5.1、6.0、4.4个百分点,在准确性方面优于其他方法,能更好地完成遥感图像目标检测任务。  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7860930  
Indexed Date:  
2025-02-23  
Institution Address:  
广西民族大学电子信息学院广西应用数学中心, 南宁, 广西 530006, 中国  
Institution Address:  
广西民族大学电子信息学院广西应用数学中心;;梧州学院, ;;广西机器视觉与智能控制重点实验室, 南宁;;梧州, ;;广西 530006;;543003  
Institution Address:  
The center for Applied Mathematics of Guangxi,School of Electronic Information,Guangxi Minzu University, Nanning, Guangxi 530006, China  
Institution Address:  
The center for Applied Mathematics of Guangxi,School of Electronic Information,Guangxi Minzu University;;Wuzhou University, ;;Guangxi Key Laboratory of Machine Vision and Intelligent Control, Nanning;;Wuzhou, ;;Guangxi 530006;;543003  
Institution:  
广西民族大学电子信息学院广西应用数学中心  
Institution:  
广西民族大学电子信息学院广西应用数学中心;;梧州学院  
Institution:  
The center for Applied Mathematics of Guangxi,School of Electronic Information,Guangxi Minzu University  
Institution:  
The center for Applied Mathematics of Guangxi,School of Electronic Information,Guangxi Minzu University;;Wuzhou University

### 619. An Improved YOLOv8 Network for Multi-Object Detection with Large-Scale

摘要: Aiming to address the challenges of low object detection precision in remote sensing images due to high background complexity and significant target scale variations, a novel model for large-scale disparate object detection in remote sensing images is proposed based on the modified YOLOv8. The model incorporates a Context Aggregation Module (CAM) with an attention mechanism in the backbone network to exploit contextual information, enabling multi-scale feature fusion for effective small object detection. The Neck network utilizes GSConv modules and employs the Slim-Neck design paradigm to enhance model robustness, making it better suited for detecting objects in high-complexity backgrounds. Furthermore, the model adopts the Wise-IoU as the loss function, incorporating a dynamic nonmonotonic focusing mechanism and a gradient gain allocation strategy to enhance the overall performance of disparate object detection. The experimental results indicate that promising performance improvements in the face of large-scale variations in remote sensing image targets. Specifically, the model achieves mAP values of 71.4% and 90.3% on the DOTA and NWPU VHR-10 datasets, respectively, representing increases of 4.4% and 3.7% compared to the original model. Compared to other typical algorithms, it also has considerable advantages in both comprehensive detection accuracy and detection speed.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0218-0014  
eISSN:  
1793-6381  
入藏号:  
WOS:001361948500008  
Indexed Date:  
2024-12-03  
Institution Address:  
Sichuan Univ Sci & Engn, Coll Automat & Informat Engn, Yibin 644000, Sichuan, Peoples R China  
Institution Address:  
Sichuan Univ Sci & Engn, Artificial Intelligence Key Lab Sichuan Prov, Yibin 644000, Sichuan, Peoples R China  
Institution Address:  
Sichuan Univ Sci & Engn, Key Lab Higher Educ Sichuan Prov Enterprise Inform, Yibin 644000, Sichuan, Peoples R China  
Institution Address:  
China West Normal Univ, Coll Comp Sci, Nanchong 637009, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 620. Lung-RetinaNet: Lung Cancer Detection Using a RetinaNet With Multi-Scale

摘要: Lung cancer is one of the terrible diseases in various countries around the globe, and timely detection of the illness is still a challenging process. The oncologists consider the blood test results and CT scans to assess the tumor, which is time-consuming and involves extra human effort. Therefore, an automated system should be developed to efficiently recognize lung tumors and assess their severity to reduce mortality. Although various researchers have proposed lung disease detection systems, the existing techniques still lack significant detection accuracy for early-stage tumors. Thus, this study proposes a novel and efficient lung tumor detector based on a RetinaNet, namely Lung-RetinaNet. A multi-scale feature fusion-based module is introduced to aggregate various network layers, simultaneously increasing the semantic information from the shallow prediction layer. Moreover, a dilated and lightweight algorithm is employed for the context module to combine contextual information with each network stage layer to improve features and effectively localize the tiny tumors. The proposed methodology attained 99.8% accuracy, 99.3% recall, 99.4% precision, 99.5% F1-score, and 0.989 Auc. We evaluated our suggested model and matched the performance with state-of-the-art DL-based methods. The outcomes show that our technique provides more substantial results than existing methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
12  
ISSN:  
2169-3536  
入藏号:  
WOS:001005704500001  
Indexed Date:  
2023-06-24  
Institution Address:  
Univ Engn & Technol UET Taxila, Dept Comp Sci, Taxila 47050, Pakistan  
Institution Address:  
King Saud Univ, Comp Sci, Riyadh, Saudi Arabia  
Institution Address:  
Georgia Univ, Comp Sci, Athens, GA USA  
Institution Address:  
Oklahoma State Univ, Dept Comp & Informat Sci, Stillwater, OK USA  
Institution Address:  
King Saud Univ, Comp Coll Vice Dean, Dept Comp Sci, Riyadh, Saudi Arabia  
Institution:  
Univ Engn & Technol UET Taxila  
Institution:  
Institution:  
Georgia Univ  
Institution:  
Institution:

### 621. A light weight multi-scale feature fusion steel surface defect detection

摘要: In industrial production, the steel surface may incur different defects owing to the influence of external factors, thereby affecting the performance of steel. With the increasing requirements for steel quality, achieving efficient detection of steel surface defects is a difficult problem that urgently needs to be solved. Traditional steel surface defect detection methods are limited by poor detection performance and slow detection speed. Therefore, a model named LMS-YOLO, based on YOLOv8, is proposed in this paper for achieving efficient steel surface defect detection. Firstly, in backbone, the light weight multi-scale mixed convolution (LMSMC) module is designed to fuse with C2f to obtain C2f\_LMSMC, so as to extract the features of different scales for fusion and achieve the light weight of the network. Meanwhile, the proposed efficient global attention mechanism was added to backbone to enhance cross dimensional information interaction and feature extraction capabilities, and to achieve a more efficient attention mechanism. In neck, using channel tuning to achieve better cross scale fusion in BiFPN. Finally, the model uses three independent decoupled heads for regression and classification, and replaces CIoU with NWD as the regression loss to enhance the effect of detecting small scale defects. The experimental results showed that LMS-YOLO achieved 81.1 mAP and 61.3 FPS on NEU-DET, 80.5 mAP and 61.3 FPS on GC10-DET, respectively. The mAP increased by 2.8 and 4.7 compared to YOLOv8, and decreased by 17.4% in floating point operations (GFLOPs) and 34.2% in parameters (Params), which indicates that the model proposed in this paper has a better comprehensive performance compared with other methods in steel surface defect detection.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
22  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
23  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001166614800001  
Indexed Date:  
2024-02-29  
Institution Address:  
Zhejiang Univ Sci & Technol, Sch Informat & Elect Engn, 318 Liuhe Rd, Hangzhou 310023, Zhejiang, Peoples R China  
Institution:

### 622. Pulmonary Nodule Classification Network Based on Multi-Scale Feature

摘要: The classification of pulmonary nodule is one of the important issues in early detection and diagnosis of lung cancer. To address the problem of information redundancy in multi-scale feature fusion and lack of discriminative feature representation in existing lung nodule classification methods, a multi-scale feature complementation and aggregate constraint (MFCAC) pulmonary nodule classification network is proposed. A multi-scale feature complementation module is proposed to learn the difference information of adjacent scale features, thereby avoiding information redundancy in the feature fusion process. Meanwhile, aggregate constraint loss is introduced into the network feature layer to achieve aggregation of similar features and improve the discriminative feature representation ability of the network. The two modules are integrated into the encoder-decoder architecture to form MFCAC, which can achieve efficient classification. Comparative experiments are conducted on the LIDC-IDRI dataset, and ablation experiments are used to analyze the contributions and effects of each component in this method. The results show that MFCAC has better performance in lung nodule classification compared to the compared algorithms.  
摘要:  
肺结节分类问题是早期肺癌检测与诊断的重要问题之一,针对现有的肺结节分类方法存在多尺度特征融合的信息冗余和缺乏判别性特征表示等问题,提出了一个基于多尺度特征互补与聚合约束(Multi-scale Feature Complementation and Aggregate Constraint, MFCAC)的肺结节分类方法,并提出了多尺度特征互补模块用于学习相邻尺度特征的差异信息,从而避免特征融合过程中的信息冗余;同时在网络特征层引入了聚合约束损失,实现对同类特征的聚集,提高网络判别性特征表示能力;将两个模块融入在编码器-解码器架构中形成MFCAC,共同作用实现高效分类。本文在LIDC-IDRI数据集上进行了对比实验,并通过消融实验分析了该方法中各组成部分的贡献和影响,结果表明,相较于对比算法,MFCAC在肺结节分类上具有更优的性能。  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-3080  
入藏号:  
CSCD:7778674  
Indexed Date:  
2024-10-12  
Institution Address:  
华东理工大学信息科学与工程学院, 上海 200237, 中国  
Institution Address:  
School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China  
Institution:  
华东理工大学信息科学与工程学院  
Institution:  
School of Information Science and Engineering, East China University of Science and Technology

### 623. DS-MSFF-Net: Dual-path self-attention multi-scale feature fusion network

摘要: Computed tomography (CT) is an important technique that is widely used in disease screening and diagnosis. In order to assist doctors in diagnosis and treatment plans, an efficient and accurate automatic image segmentation technology is urgently needed. CT images of different lesions always have problems such as different resolutions, different numbers of lesions, and inconspicuous contrast between lesions and background areas, which brings considerable challenges to the automated segmentation process. To this end, we propose a dual-path self-attention multi-scale feature fusion network (DS-MSFF-Net) that fuses self-attention mechanism and dilated convolution. It is worth noting that this network includes two parallel branch paths, which enables it to extract long-range semantic feature information effectively while extracting detailed feature information of CT images. Additionally, a novel feature extraction module is designed to focus limited learning resources on low-resolution high-order semantic feature maps, which can improve the segmentation accuracy without significant additional computational overhead. We extensively evaluate our method on the LIDC-IDRI lung nodule segmentation dataset and the LiTS2017 liver segmentation dataset, which outperforms other recent state-of-the-art methods on various CT image segmentation tasks.  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:001194586400002  
Indexed Date:  
2024-04-07  
Institution Address:  
Southwest Univ Sci & Technol, Sch Informat Engn, Mianyang 621010, Peoples R China  
Institution Address:  
Southwest Univ Sci & Technol, Sch Comp Sci & Technol, Mianyang 621010, Peoples R China  
Institution Address:  
Mianyang Cent Hosp, Radiol Dept, Mianyang 621010, Peoples R China  
Institution:  
Institution:  
Institution:  
Mianyang Cent Hosp

### 624. Mixed-scale cross-modal fusion network for referring image segmentation

摘要: Referring image segmentation aims to segment the target by a given language expression. Recently, the bottom-up fusion network utilizes language features to highlight the most relevant regions during the visual encoder stage. However, it is not comprehensive that establish only the relationship between pixels and words. To alleviate this problem, we propose a mixed-scale cross-modal fusion method that widens the interaction between vision and language. Specially, at each stage, pyramid pooling is used to augment visual perception and improve the interaction between visual and linguistic features, thereby highlighting relevant regions in the visual data. Additionally, we employ a simple multi-scale feature fusion module to effectively combine multi-scale aligned features. Experiments conducted on Standard RIS benchmarks demonstrate that the proposed method achieves favorable performance against state-of-the-art approaches. Moreover, we conducted experiments on different visual backbones respectively, and the proposed method yielded better and significantly improved performance results.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:001349667100001  
Indexed Date:  
2024-11-14  
Institution Address:  
Xidian Univ, Sch Artificial Intelligence, Xian 710071, Peoples R China  
Institution Address:  
Xidian Univ, Guangzhou Inst Technol, Guangzhou 510700, Peoples R China  
Institution Address:  
Pazhou Lab Huangpu, Guangzhou 510555, Peoples R China  
Institution Address:  
Shanxi Datong Univ, Datong 037009, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 625. Occluded Pedestrian Detection Based on Multi-Scale Context Information

摘要: Pedestrian detection in occluded scenes has always been a thorny problem in computer vision. In this case, due to the large difference in scale of occluded pedestrians and low visibility, it usually brings great challenges to detection. To solve this problem, this paper proposes a model structure for pedestrian occlusion detection, which improves the pedestrian detection method based on anchor-free. First, a structure for extracting multi-scale context information is designed. By cascading multiple convolutional layers with different dilation rates, using dense connections to achieve multi-scale feature sharing, the context information of each region is extracted to solve the occlusion problem. In addition, in order to improve the robustness of features, the multi-scale feature fusion is adaptive adjusted using the channel attention mechanism. Experimental results show that this method achieves 41.73% of MR~(-2) on the occlusion subset of Caltech pedestrian dataset, which is better than other contrast detectors.  
摘要:  
在遮挡场景下的行人检测一直是计算机视觉中的一个棘手问题,由于被遮挡的行人尺度差异大,可见率低,通常会给检测带来极大的挑战。针对这一问题,提出了一种针对行人遮挡检测的模型结构,对基于anchor-free的行人检测方法进行改进。设计了一种提取多尺度上下文信息的结构,通过级联多个不同扩张率的卷积层,使用密集连接实现多尺度特征共享,提取各个区域的上下文信息来解决遮挡问题。此外,为了提高特征的可分辨性,使用通道注意力机制对多尺度特征融合进行自适应的调整。实验结果表明,该方法在Caltech行人数据集的遮挡子集上实现了41.73%的MR~(-2),性能优于其他检测算法。  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7227026  
Indexed Date:  
2022-08-26  
Institution Address:  
上海海事大学信息工程学院, 上海 201306, 中国  
Institution Address:  
College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China  
Institution:  
上海海事大学信息工程学院  
Institution:  
College of Information Engineering, Shanghai Maritime University

### 626. Lightweight Context Awareness and Feature Enhancement for Anchor-Free

摘要: Optical remote-sensing image target detection holds significant research significance in various domains, including disaster relief, ecological environment protection, and military surveillance. However, since remote-sensing images have multiscale targets, complex backgrounds, and many small targets, the performance of the existing network models in remote-sensing image target detection cannot reach what we expect. In addition, we note that current networks use complex computational mechanisms that make the models time-consuming, which hinders their practicability in remote-sensing target detection scenarios. In response to this challenge, we propose an anchor-free and efficient one-stage target detection method for optical remote-sensing images. First, we propose the lightweight context-aware module GSelf-Attention, injected into the feature fusion network from top-to-bottom and bottom-to-top to enhance the feature information interaction. Second, we proposed that ELAN-RSN uses an optimized residual shrinkage network (RSN) to eliminate background noise and conflicting information in the multiscale feature fusion. Finally, we introduce the decoupled head fused with SPDConv to enhance the detection accuracy of small target objects further. The performance of the proposed algorithm is compared with that of other advanced methods on DIOR and RSOD datasets. The experimental results show that the proposed algorithm significantly improves object detection accuracy while ensuring detection efficiency and has high robustness. The code is available at https://github.com/FF-codeHouse/Object-Detection/tree/remote-sensing.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001245641800079  
Indexed Date:  
2024-06-17  
Institution Address:  
Inner Mongolia Univ Sci & Technol, Sch Digital & Intelligence Ind, Baotou 014010, Peoples R China  
Institution:

### 627. Intelligent detection of fine cracks on sleepers based on improved

摘要: The detection of cracks in sleepers is important in ensuring safe operation of trains and the crack width is a key index in this process. The traditional YOLOv5 model can not achieve satisfactory detection results in the face of large changes in the scale of the target and interference of background information in the detection process. In this paper, a method to detect and measure the fine crack of sleeper based on improved YOLOv5 model of cascade fusion and edge fitting algorithm is proposed to solve these problems. An adaptive multi-scale feature fusion technique is proposed to achieve multi-scale feature fusion efficiently and solve the problem of large changes in the scale of the target. An improved method of feature recombination of attention mechanism is used to solve the interference problem of background information. On this basis, a method based on cascade fusion is proposed to fuse the above improvements and reconstruct the structure of the YOLOv5 model. Moreover, a fitting algorithm of the crack edge is proposed to realize the precise measurement of the maximum width of the fine crack on sleepers. Finally, the process of crack detection and width measurement of sleeper is developed and the performance of the method proposed in this paper is verified by experiments through self-made data sets. The experimental results show that the precision of crack detection is increased from 84.08% to 96.86%, and the recall is increased from 78.13% to 96.46%. The increases are 12.78% and 18.33% respectively. At the same time, the average detection time is reduced from 15.23 ms to 10.04 ms. The measurement accuracy of the maximum crack width is 0.05 mm.  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0954-4097  
eISSN:  
2041-3017  
入藏号:  
WOS:000948621400001  
Indexed Date:  
2023-03-27  
Institution Address:  
China Acad Railway Sci Corp Ltd, Railway Engn Res Inst, Beijing 100081, Peoples R China  
Institution Address:  
Soochow Univ, Sch Mech & Elect Engn, Suzhou, Peoples R China  
Institution:  
China Acad Railway Sci Corp Ltd  
Institution:

### 628. TMS-GAN: A Twofold Multi-Scale Generative Adversarial Network for Single

摘要: In recent years, learning-based single image dehazing networks have been comprehensively developed. However, performance improvement is limited due to domain shift between trained synthetic hazy images and untrained real-world hazy images. To alleviate this issue, this paper proposes a real-world dehazing targeted training scheme which nearly realizes paired real-world data training. As a result, a Twofold Multi-scale Generative Adversarial Network (TMS-GAN) consisting of a Haze-generation GAN (HgGAN) and a Haze-removal GAN (HrGAN) is designed. HgGAN attributes real haze properties to synthetic images and HrGAN removes haze from both synthetic and generated fake realistic data under supervision. Thus, the proposed method can better adapt to real-world image dehazing using this cooperative training scheme. Meanwhile, several structural advances of TMS-GAN also improve dehazing performance. Specifically, a haze residual map based on atmospheric scattering model is deduced in HgGAN for fake realistic data generation. The dual-branch generator in HrGAN draws attention to detail restoration by one branch along with another color-branch. A plug-and-play Multi-attention Progressive Fusion Module (MAPFM) is proposed and inserted in both HgGAN and HrGAN. MAPFM incorporates multi-attention mechanism to guide multi-scale feature fusion in a progressive manner, in which Adjacency-attention Block (AAB) can capture contributing features of each level and Self-attention Block (SAB) can establish non-local dependency of feature fusion. Experiments on mainstream benchmarks show that the proposed framework is superior especially on real-world hazy images among single image dehazing methods.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
43  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
44  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:000790830300024  
Indexed Date:  
2022-05-15  
Institution Address:  
East China Univ Sci & Technol, Sch Informat Sci & Engn, Shanghai 200237, Peoples R China  
Institution Address:  
Shanghai Normal Univ, Coll Informat & Elect Engn, Shanghai 200237, Peoples R China  
Institution:  
Institution:

### 629. Panoptic segmentation algorithm based on grouped convolution for feature

摘要: Aiming at the problem that the computing of the image panoptic segmentation task is not fast enough for the existing network structures in practical applications,a panoptic segmentation algorithm based on grouped convolution for feature fusion was proposed.Firstly,through the bottom-up method,the classic Residual Network structure(ResNet)was selected for feature extraction,and the multi-scale feature fusion of semantic segmentation and instance segmentation was performed on the extracted features by using the Atrous convolutional Spatial Pyramid Pooling operation (ASPP) with different expansion rates.Secondly,a single-channel grouped convolution upsampling method was proposed to integrate the semantics and instance features for performing upsampling feature fusion to a specified size.Finally,a more refined panoptic segmentation output result was obtained by performing loss function on semantic branch,instance branch and instance center point respectively.The model was compared with Attention-guided Unified Network for panoptic segmentation(AUNet),Panoptic Feature Pyramid Network(Panoptic FPN),Single-shot instance Segmentation with Affinity Pyramid(SSAP),Unified Panoptic Segmentation Network(UPSNet),Panoptic-DeepLab and other methods on CityScapes dataset.Compared with the Panoptic-DeepLab model,which is the best-performing model in the comparison models,with the decoding network parameters reduced significantly,the proposed model has the Panoptic Quality(PQ)of 0.565,with a slight decrease of 0.003,and the segmentation qualities of objects such as buildings,trains,bicycles were improved by 0.3 - 5.5,the Average Precision(AP)and the Average Precision with target IoU(Intersection over Union)threshold over 50%(AP50) were improved by 0.002 and 0.014 respectively,and the mean IoU(mIoU)value was increased by 0.06.It can be seen that the proposed method improves the speed of image panoptic segmentation,has good accuracy in the three indexes of PQ,AP and mIoU,and can effectively complete the panoptic segmentation tasks.  
摘要:  
针对图像全景分割任务对于实践应用中现有网络结构运算不够快速的问题,提出一种基于分组卷积进行特征融合的全景分割算法。首先,通过自底向上的方式选择经典残差网络结构(ResNet)进行特征提取,并采用不同扩张率的空洞卷积空间金字塔池化操作(ASPP)对提取到的特征进行语义分割与实例分割的多尺度特征融合;然后,通过提出一种单路分组卷积上采样方法,整合语义与实例特征进行上采样特征融合至指定大小;最后,通过对语义分支、实例分支以及实例中心点这三个分支进行损失函数运算以得到更加精细的全景分割输出结果。该模型在CityScapes数据集上与注意力引导的联合全景分割网络(AUNet)、全景特征金字塔网络(Panoptic FPN)、亲和金字塔单阶段实例分割算法(SSAP)、联合全景分割网络(UPSNet)、Panoptic-DeepLab等方法进行了实验对比。实验结果表明,与对比方法中表现最好的Panoptic-DeepLab模型相比,所提模型在极大减少了解码网络参数量的情况下,全景分割质量(PQ)值为0.565,仅下降了0.003,在建筑物、火车、自行车等物体的分割质量上有0.3~5.5的提升,平均精确率(AP)、目标IoU阈值超过50%的平均精确率(AP50)分别提升了0.002与0.014,平均交并比(mIoU)值提升了0.06。可见该方法能提升图像全景分割速度,在PG、AP、mIoU三个指标上均有较好的精度,可以有效地完成全景分割任务。  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
5  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
7  
ISSN:  
1001-9081  
入藏号:  
CSCD:7011186  
Indexed Date:  
2021-10-09  
Institution Address:  
中国民航大学计算机科学与技术学院;;中国民航大学信息网络中心, ;;, ;;, 天津;;天津 300300;;300300, 中国  
Institution Address:  
中国民航大学计算机科学与技术学院, 天津 300300, 中国  
Institution Address:  
College of Computer Science and Technology,Civil Aviation University of China;;Information Network Center,Civil Aviation University of China, ;;, ;;, Tianjin;;Tianjin 300300;;300300  
Institution Address:  
College of Computer Science and Technology,Civil Aviation University of China, Tianjin 300300, China  
Institution:  
中国民航大学计算机科学与技术学院;;中国民航大学信息网络中心  
Institution:  
中国民航大学计算机科学与技术学院  
Institution:  
College of Computer Science and Technology,Civil Aviation University of China;;Information Network Center,Civil Aviation University of China  
Institution:  
College of Computer Science and Technology,Civil Aviation University of China

### 630. A novel approach for underwater fish segmentation in complex scenes

摘要: Underwater segmentation technology achieves effective monitoring of fish biological information through accurate identification of fish species and precise estimation of their quantities. It serves as an effective approach to enhance the informatization level of aquaculture and promote intelligent management in fisheries. However, the complex and diverse underwater environment, coupled with poor visibility, results in blurry and lower-quality underwater fish images. The current image segmentation methods, when applied to fish segmentation, exhibit low accuracy and inadequate generalization capabilities. This paper proposes an image segmentation model based on Atrous Spatial Pyramid Pooling (ASPP) to address these challenges. The model aims to improve fish feature extraction and enhance the segmentation precision of fish images in complex underwater environments, thereby enhancing the accuracy and generalization capabilities of existing fish segmentation models. First, a multiscale feature extraction module (triangular atrous spatial pyramid multifeature fusion) based on dilated convolutional spatial pyramid pooling, which enhances the extraction of high-level semantic features of images through the triangular combination of multilayer dilated convolutional pyramid pooling and the adaptive channel attention module, is proposed. Second, a spatial attention module based on strip pooling (atrous strip pooling) is proposed, which further expands the receptive field of the attention mechanism by combining different expansion rates, enhances the correlations between pixels, and effectively captures spatial information. Finally, a decoder module based on multilayer semantic feature fusion is proposed. Through the processing and fusion of medium-, low-, and high-level semantic features, the model understands image content and performs accurate pixel-level segmentation. The proposed model is evaluated using a VOC-compliant dataset created from underwater fish images and validated against public and specific underwater fish datasets. The results demonstrate the successful application of the feature extraction and feature fusion modules in underwater fish image segmentation, achieving an average Intersection over Union (MIoU) of 85.49% in segmentation tasks. Compared to conventional segmentation models, the proposed model shows significant improvements, with an average increase of 3.8% in MIoU and 2.5% in balanced F-score (F1-score). The accurate segmentation of underwater fish images and extraction of vital biological information by this model provide a solid foundation for intelligent monitoring, including fish length measurement, weight estimation, and analysis of growth and health status. Moreover, the model offers a scientific framework for decision-making in aquaculture, driving advancements in precision and intelligent management practices in the field.  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
2  
ISSN:  
0967-6120  
eISSN:  
1573-143X  
入藏号:  
WOS:001170784500002  
Indexed Date:  
2024-03-19  
Institution Address:  
Shandong Technol & Business Univ, Sch Comp Sci & Technol, Yantai 264005, Shandong, Peoples R China  
Institution:

### 631. The visual motion blur elimination method for silicon nitride bearing

摘要: The visual motion blur imaging for the feature recognition process of silicon nitride bearing roller fissures is a pathological problem. This is solved by proposing squeeze-and-excitation asymmetric fusion of multi-scale features with high-frequency loss attention coupled U-Net (MHU-Net). The visual motion blur elimination of fissure features on silicon nitride bearing rollers is achieved. In the deblurring model, the multi-scale feature information on silicon nitride bearing roller fissures is blocked and there is weak correlation between channels. A design for an asymmetric fusion multi-scale feature module under the channel information compression-excitation mode is proposed. It successfully balances the channel information from different scales while integrating multi-scale features in image fusion. The high-frequency region of fissure features on silicon nitride bearing rollers is analyzed. Around the high-frequency feature loss in the multi-frequency domain of images combined with spatial feature loss, a multi-frequency band high-frequency loss attention module is built. Then, the complete structural details of silicon nitride bearing roller fissures are obtained. The proposed algorithm achieves a peak signal-to-noise ratio of 27.58 and a structural similarity of 0.847 on our self-made silicon nitride defect motion dataset. The visual motion blur of fissure features is noticeably eliminated. The restored image exhibits complete details in the feature structures and overall region smoothness. (c) 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2158-3226  
入藏号:  
WOS:001234684800002  
Indexed Date:  
2024-06-04  
Institution Address:  
Jingdezhen Ceram Univ, Sch Informat Engn, Jingdezhen 333403, Peoples R China  
Institution Address:  
Jingdezhen Ceram Univ, Sch Mech & Elect Engn, Jingdezhen 333403, Peoples R China  
Institution Address:  
Lab Ceram Mat Proc Technol Engn, Jingdezhen 333403, Jiangxi, Peoples R China  
Institution:  
Institution:  
Institution:  
Lab Ceram Mat Proc Technol Engn

### 632. MS<SUP>2</SUP>A<SUP>2</SUP>Net: Multiscale Self-Attention Aggregation

摘要: Few-shot aerial imagery segmentation refers to the task of segmenting specific objects in scenes that have not been encountered during training with a small amount of annotated data for reference. However, most existing few-shot segmentation algorithms are primarily designed for natural images, and there is still a lack of exploration in the context of remote sensing aerial imagery. In this article, we propose a novel multiscale self attention aggregation network (MS(2)A(2)Net), dubbed MS(2)A(2)Net, to address the challenge of few-shot aerial image segmentation in terms of scarce data and network architecture. Specifically, we first incorporate the designed asymmetric momentum contrastive learning (AMCL) into the pre-training stage, to improve the representation capability of the backbone without the expensive labeled data. Then the frozen encoder is transferred to the downstream few-shot segmentation task as the feature embedding. In terms of network architecture, we design self-attention aggregation in multiscale feature fusion, to construct the dual correlation of foreground and background between support and query features at the pixel level. Besides, the coordinate attention is designed to rearrange the distribution of feature importance in both horizontal and vertical spatial order perspectives, which facilitates adaptive fusion with the multiscale features. To verify the availability of the proposed MS(2)A(2)Net, we also reconstructed two novel datasets dedicated to few-shot aerial image segmentation, called DLRSD-4(i) and iSAID-4(i). The experimental results show that our approach MS(2)A(2)Net is superior in three few-shot benchmark aerial imagery segmentation datasets, which achieves competitive segmentation performance. Extensive ablation experiments also reflect the effectiveness and scalability of the proposed components and overall network architecture.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
8  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001129141000017  
Indexed Date:  
2024-03-13  
Institution Address:  
Xidian Univ, Key Lab Collaborat Intelligence Syst, Minist Educ, Xian, Peoples R China  
Institution:

### 633. I-CenterNet: Road infrared target detection based on improved CenterNet

摘要: Infrared target detection has strong anti-interference ability, long working distance and can work day and night. So it is widely used in military security and transportation fields, and infrared road object detection is critical in traffic checkpoints and autonomous driving. However, the target scale in infrared images changes greatly, small targets are difficult to detect, the poor image quality and low signal-to-noise ratio are still huge challenges in infrared target detection. This paper proposes an improved infrared target detection model I-CenterNet based on the anchor-free model CenterNet. The EfficientNetV2 with the channel attention mechanism is used instead of the traditional structure as the backbone network to enhance feature extraction. In order to reduce the noise of the input infrared image, Dilated-Residual U-net (DRUNet) is used. Meanwhile, feature pyramid and Sub-Pixel are combined for multi-scale feature fusion. Data enhancement is implemented to improve model performance. The experimental results show that the average detection accuracy of this model on the Flir infrared data set is 87.9%, and the average detection speed reaches 14.2 frames/s.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1751-9659  
eISSN:  
1751-9667  
入藏号:  
WOS:000851504400001  
Indexed Date:  
2022-09-17  
Institution Address:  
Xinjiang Univ, Sch Software, Urumqi 830046, Peoples R China  
Institution Address:  
Key Lab Signal Detect & Proc Xinjiang Uygur Auton, Urumqi, Peoples R China  
Institution Address:  
Xinjiang Univ, Software Engn Lab, Urumqi, Peoples R China  
Institution Address:  
Lianhai Chuangzhi Informat Technol Co Ltd, Urumqi, Xinjiang Uygur, Peoples R China  
Institution:  
Institution:  
Key Lab Signal Detect & Proc Xinjiang Uygur Auton  
Institution:  
Institution:  
Lianhai Chuangzhi Informat Technol Co Ltd

### 634. MFF: An effective method of solving the ill regions in stereo matching

摘要: In the current stereo matching field, the accuracy of the derived disparity map is highly dependent on the processing capability of ill regions. Fortunately, we find that the use of local information will eliminate the negative effects associated with ill regions. As a result of the above discovery, we propose the Concatenated Dilated Convolution (CDC) block and the Multi-scale Feature Fusion module (MFF), which are capable of effectively extracting regional context by increasing the receptive field in parallel and channel-wise ways. The CDC block can expand the receptive field by applying multiple dilated convolutions at different dilation rates in parallel to enhance the smoothness of the feature map. By constructing parallel CDC blocks in a multiple dilated manner, the MFF module can improve the smoothness of the feature map. In addition, to control the number of parameters in the MFF network, a high-performance channel-distribution algorithm is proposed, capable of adjusting the weights of each module and convolution in an adaptive manner while reducing the number of parameters. Extensive experiments have demonstrated that MFF and CDC can effectively improve the performance of ill areas and networks with a minimal number of parameters.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1751-9632  
eISSN:  
1751-9640  
入藏号:  
WOS:000975273700001  
Indexed Date:  
2023-05-09  
Institution Address:  
Guizhou Univ, Coll Comp Sci & Technol, Guiyang, Peoples R China  
Institution Address:  
Guizhou Univ Finance & Econ, Sch Informat, Guiyang, Peoples R China  
Institution Address:  
Peking Univ, Sch Elect & Comp Engn, Shenzhen Grad Sch, Shenzhen, Peoples R China  
Institution:  
Institution:  
Institution:

### 635. A Low-Cost and Lightweight Real-Time Object-Detection Method Based on

摘要: Accurate detection of transportation objects is pivotal for enhancing driving safety and operational efficiency. In the rapidly evolving domain of transportation systems, the utilization of unmanned aerial vehicles (UAVs) for low-altitude detection, leveraging remotely-sensed images and videos, has become increasingly vital. Addressing the growing demands for robust, real-time object-detection capabilities, this study introduces a lightweight, memory-efficient model specifically engineered for the constrained computational and power resources of UAV-embedded platforms. Incorporating the FasterNet-16 backbone, the model significantly enhances feature-processing efficiency, which is essential for real-time applications across diverse UAV operations. A novel multi-scale feature-fusion technique is employed to improve feature utilization while maintaining a compact architecture through passive integration methods. Extensive performance evaluations across various embedded platforms have demonstrated the model's superior capabilities and robustness in real-time operations, thereby markedly advancing UAV deployment in crucial remote-sensing tasks and improving productivity and safety across multiple domains.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
eISSN:  
2072-4292  
入藏号:  
WOS:001332817300001  
Indexed Date:  
2024-10-21  
Institution Address:  
Xidian Hangzhou Inst Technol, Hangzhou 311231, Peoples R China  
Institution Address:  
Univ Tunku Abdul Rahman, Lee Kong Chian Fac Engn & Sci, Kajang 43000, Malaysia  
Institution Address:  
Xidian Guangzhou Inst Technol, Guangzhou 510555, Peoples R China  
Institution:  
Xidian Hangzhou Inst Technol  
Institution:  
Institution:  
Xidian Guangzhou Inst Technol

### 636. Multi-Scale Learnable Gabor Transform for Pedestrian Trajectory

摘要: Trajectory prediction is an important task in autonomous driving and monitoring systems. Most of the existing methods pay little attention to the rapidly changing trajectory information, but how to effectively solve this problem is crucial to ensure pedestrian safety. The Gabor transform has inherent advantages for capturing instantaneously changing information. Therefore, for the first time, we introduce the Gabor transformation idea into pedestrian trajectory prediction and propose the Multi-scale Learnable Gabor Transform Network (MlgtNet), which establishes global and local contextual relationships from multi-dimensional and multi-scale perspectives. The network first uses the Multi-scale Feature Dimension Enhancement Module (MFDEM) ascending dimension trajectory sequence, and uses the Multi-scale Gabor Convolution Module (MGCM) to guide the model to establish the dependence of different distances from different dimensions to model the interrelationship between global/local features at different scales and different step sizes. Finally, the Feature Fusion Module (FFM) processes the multimodal information and fuses it with the multi-scale trajectory features to obtain the trajectory prediction representation in different visual fields. The representation results are then used for secondary fusion to obtain the global prediction results. Experimental results show that MlgtNet achieves state-of-the-art performance with its lightweight model size on the vast majority of widely used trajectory prediction datasets from different perspectives.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1524-9050  
eISSN:  
1558-0016  
入藏号:  
WOS:001271447100001  
Indexed Date:  
2024-07-26  
Institution Address:  
Northeastern Univ, Sch Informat Sci & Engn, Shenyang 110819, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Sch Comp Sci & Engn, Guangzhou 510275, Peoples R China  
Institution:  
Institution:

### 637. CapsLoc3D: Point Cloud Retrieval for Large-Scale Place Recognition Based

摘要: Point cloud-based place recognition can be used for global localization in large-scale scenes and loop-closure detection in simultaneous localization and mapping (SLAM) systems in the absence of GPS. Current learning-based approaches aim to extract global and local features from 3D point clouds to encode them as descriptors for point cloud retrieval. The key problems are that the occlusion of point clouds by dynamic objects in the scene affects the point cloud structure, a single perceptual field of the network cannot adequately extract point cloud features, and the correlation between features is not fully utilized. To overcome this, we propose a novel network called CapsLoc3D. We first use the static point cloud generation module to remove the occlusion effects of dynamic objects, and then obtain the point cloud descriptors by processing with the CapsLoc3D network which contains the point spatial transformation module, multi-scale feature fusion module, Capsnet module and a GeM Pooling layer. After validation using the Oxford RobotCar, KITTI, and NEU datasets, experiments show that our method performs better and also has good generalization performance and computational efficiency compared with current state-of-the-art algorithms.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1524-9050  
eISSN:  
1558-0016  
入藏号:  
WOS:001167302500001  
Indexed Date:  
2024-03-02  
Institution Address:  
Northeastern Univ, Coll Informat Sci & Engn, Shenyang 110819, Peoples R China  
Institution Address:  
Ulster Univ, Sch Comp Engn & Intelligent Syst, Coleraine BT52 1SA, North Ireland  
Institution:  
Institution:

### 638. ELGS-YOLO: small object detection method of power switchgear images

摘要: To address the challenges of difficult detection of small targets and limited computational resources prevalent in power switchgear. This paper presents an enhanced switchgear detection model, ELGS-YOLOv8, which addresses the limitations of existing models. First, ELGS-YOLOv8 employs a lightweight Efficient Mutil Scal Conv (EMSC) in the backbone network. This module reduces the number of parameters and computation of the model while improving the extraction of feature information. In addition, the local-global information extraction and fusion module (LGEF) for small targets has been proposed to further improve the detection capability of the indicator. The module integrates local information extraction module(LIEM)and local-global information fusion module (LGFM) at the neck, which are used to improve the extraction and fusion of feature information, respectively. Finally, a lightweight detection head based on the Separated and Enhancement Attention Module (LSEAM) is designed to reduce the model complexity while further improve the model's detection performance in complex scenes. Experimental results show that the improved ELGS-YOLOv8 model achieves 90.4% in mAP@0.5, with a model size reduction of 15% and an increase in APs by 7.7%, meeting the real-time and accuracy requirements for on-site deployment. Therefore, the improved strategy provides a more reliable and efficient solution for switchgear detection in substation inspections.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2631-8695  
入藏号:  
WOS:001432871900001  
Indexed Date:  
2025-03-05  
Institution Address:  
Zhengzhou Univ Light Ind, Coll Elect & Informat Engn, Zhengzhou 450000, Peoples R China  
Institution:

### 639. An Infrared Image Defect Detection Method for Steel Based on Regularized

摘要: Steel surfaces often display intricate texture patterns that can resemble defects, posing a challenge in accurately identifying actual defects. Therefore, it is crucial to develop a highly robust defect detection model. This study proposes a defect detection method for steel infrared images based on a Regularized YOLO framework. Firstly, the Coordinate Attention (CA) is embedded within the C2F framework, utilizing a lightweight attention module to enhance the feature extraction capability of the backbone network. Secondly, the neck part design incorporates the Bi-directional Feature Pyramid Network (BiFPN) for weighted fusion of multi-scale feature maps. This creates a model called BiFPN-Concat, which enhances feature fusion capability. Finally, the loss function of the model is regularized to improve the generalization performance of the model. The experimental results indicate that the model has only 3.03 M parameters, yet achieves a mAP@0.5 of 80.77% on the NEU-DET dataset and 99.38% on the ECTI dataset. This represents an improvement of 2.3% and 1.6% over the baseline model, respectively. This method is well-suited for industrial detection applications involving non-destructive testing of steel using infrared imagery.  
Web of Science 核心合集中的 "被引频次":  
9  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
9  
eISSN:  
1424-8220  
入藏号:  
WOS:001183041600001  
PubMed ID:  
38475212  
Indexed Date:  
2024-03-30  
Institution Address:  
Kunming Univ Sci & Technol, Fac Informat Engn & Automat, Kunming 650500, Peoples R China  
Institution:

### 640. Advancing Rice Grain Impurity Segmentation with an Enhanced SegFormer

摘要: During the rice harvesting process, severe occlusion and adhesion exist among multiple targets, such as rice, straw, and leaves, making it difficult to accurately distinguish between rice grains and impurities. To address the current challenges, a lightweight semantic segmentation algorithm for impurities based on an improved SegFormer network is proposed. To make full use of the extracted features, the decoder was redesigned. First, the Feature Pyramid Network (FPN) was introduced to optimize the structure, selectively fusing the high-level semantic features and low-level texture features generated by the encoder. Secondly, a Part Large Kernel Attention (Part-LKA) module was designed and introduced after feature fusion to help the model focus on key regions, simplifying the model and accelerating computation. Finally, to compensate for the lack of spatial interaction capabilities, Bottleneck Recursive Gated Convolution (B-gnConv) was introduced to achieve effective segmentation of rice grains and impurities. Compared with the original model, the improved model's pixel accuracy (PA) and F1 score increased by 1.6% and 3.1%, respectively. This provides a valuable algorithmic reference for designing a real-time impurity rate monitoring system for rice combine harvesters.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1099-4300  
入藏号:  
WOS:001405343000001  
PubMed ID:  
39851692  
Indexed Date:  
2025-02-01  
Institution Address:  
Jiangsu Univ Sci & Technol, Sch Automat, Zhenjiang 212100, Peoples R China  
Institution Address:  
Jiangsu Univ Sci & Technol, Sch Comp, Zhenjiang 212100, Peoples R China  
Institution Address:  
Suzhou Univ Sci & Technol, Coll Mat Sci & Engn, Suzhou 215011, Peoples R China  
Institution:  
Institution:  
Institution:

### 641. Detail texture detection based on Yolov4-tiny combined with attention

摘要: Aero-engine blades crack detection is one of the important tasks in daily ground maintenance, crack is a kind of texture feature, due to the random distribution, irregular shape and vague characteristics, which is still a challenging task to realize automatic detection in working environment. A detection model based on the Yolov4-tiny is proposed that is universal and focuses more on the characteristics of cracks, and it is implemented in embedded device. First, in order to distinguish the cracks and noises, an improved attention module is introduced into the backbone of Yolov4-tiny to enhance the model's capability to focus on crack areas; second, in order to improve the effect of multi-scale feature fusion, the bicubic interpolation is implemented in upsampling module; finally, in order to solve the redundant detection results of bounding-boxes in crack areas, the optimized non-maximum suppression method is proposed to make the detection results better corresponding to the groundTruth. The robustness of proposed detection model was demonstrated by evaluating varying lighting and noise images. The average precision on integrated datasets is 81.6%, which outperforms the original Yolov4-tiny by an increase of 12.3%.  
在 BIOSIS Citation Index 中的被引频次:  
5  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
17  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
18  
ISSN:  
1751-9659  
eISSN:  
1751-9667  
入藏号:  
WOS:000640656500001  
Indexed Date:  
2021-05-02  
Institution Address:  
Northwestern Polytech Univ, Inst Unmanned Syst Res, Xian, Peoples R China  
Institution:

### 642. A lightweight bladder tumor segmentation method based on attention

摘要: In the endoscopic images of bladder, accurate segmentation of different grade bladder tumor from blurred boundary regions and highly variable shapes is of great significance for doctors' diagnosis and patients' later treatment. We propose a nested attentional feature fusion segmentation network (NAFF-Net) based on the encoder-decoder structure formed by the combination of weighted pyramid pooling module (WPPM) and nested attentional feature fusion (NAFF). Among them, WPPM applies the cascade of atrous convolution to enhance the overall perceptual fieldwhile introducing adaptiveweights to optimize multi-scale feature extraction, NAFF integrates deep semantic information into shallow feature maps, effectively focusing on edge and detail information in bladder tumor images. Additionally, a weighted mixed loss function is constructed to alleviate the impact of imbalance between positive and negative sample distribution on segmentation accuracy. Experiments illustrate the proposed NAFF-Net achieves better segmentation results compared to other mainstream models, with aMIoU of 84.05%, MPrecision of 91.52%, MRecall of 90.81%, and F1-score of 91.16%, and also achieves good results on the public datasets Kvasir-SEG and CVC-ClinicDB. Compared to other models, NAFF-Net has a smaller number of parameters, which is a significant advantage in model deployment.  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0140-0118  
eISSN:  
1741-0444  
入藏号:  
WOS:001154262400002  
PubMed ID:  
38308022  
Indexed Date:  
2024-02-11  
Institution Address:  
Guangdong Univ Technol, Sch Automat, Guangzhou 510006, Peoples R China  
Institution Address:  
Nanchang Univ, Affiliated Hosp 1, Dept Urol, Nanchang 330006, Peoples R China  
Institution:  
Institution:

### 643. Research on a Traffic Sign Recognition Method under Small Sample

摘要: Traffic signs are updated quickly, and there image acquisition and labeling work requires a lot of manpower and material resources, so it is difficult to provide a large number of training samples for high-precision recognition. Aiming at this problem, a traffic sign recognition method based on FSOD (few-shot object learning) is proposed. This method adjusts the backbone network of the original model and introduces dropout, which improves the detection accuracy and reduces the risk of overfitting. Secondly, an RPN (region proposal network) with improved attention mechanism is proposed to generate more accurate target candidate boxes by selectively enhancing some features. Finally, the FPN (feature pyramid network) is introduced for multi-scale feature extraction, and the feature map with higher semantic information but lower resolution is merged with the feature map with higher resolution but weaker semantic information, which further improves the detection accuracy. Compared with the baseline model, the improved algorithm improves the 5-way 3-shot and 5-way 5-shot tasks by 4.27% and 1.64%, respectively. We apply the model structure to the PASCAL VOC dataset. The results show that this method is superior to some current few-shot object detection algorithms.  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
1424-8220  
入藏号:  
WOS:001005536500001  
PubMed ID:  
37299816  
Indexed Date:  
2023-06-25  
Institution Address:  
Xinjiang Univ, Coll Informat Sci & Engn, Urumqi 830017, Peoples R China  
Institution Address:  
Xinjiang Univ, Key Lab Multilingual Informat Technol Xinjiang Uyg, Urumqi 830017, Peoples R China  
Institution:  
Institution:

### 644. An enhancement model based on dense atrous and inception convolution for

摘要: The goal of semantic segmentation is to classify each pixel in the image, so as to segment out the specific contour of the target. Most previous semantic segmentation models cannot generate enough semantic information for each pixel to understand the content of complex scenes. In this paper, we propose a novel semantic segmentation model Ince-DResAsppNet based on dense convoluted separation convolution. Unlike the previous model, our model revolves around reducing semantic information loss and enhancing detailed information. In the feature extraction part of the model, the idea of Dense and Ince is introduced to expand the number of channels on the basis of feature reuse. In the feature fusion part, Dense and Atrous's idea of dense dilated based on coprime factors is introduced, combined with multi-scale feature information to expand the receptive field and collect more dense pixels. Experiments conducted on the dataset PASCAL VOC 2012 and the CityScapes dataset show that our method performs better than the existing semantic segmentation model. Our model achieves 83.3% and 78.1% segmentation accuracy on the mIoU indicator, which surpasses many classical semantic segmentation models.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
10  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
10  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000815559900001  
Indexed Date:  
2022-07-01  
Institution Address:  
Beijing Jiaotong Univ, Sch Comp & Informat Technol, Beijing, Peoples R China  
Institution Address:  
Simon Fraser Univ, Sch Comp Sci, Vancouver, BC, Canada  
Institution:  
Institution:

### 645. MSCD-YOLO: A Lightweight Dense Pedestrian Detection Model with

摘要: Pedestrian detection is widely used in real-time surveillance, urban traffic, and other fields. As a crucial direction in pedestrian detection, dense pedestrian detection still faces many unresolved challenges. Existing methods suffer from low detection accuracy, high miss rates, large model parameters, and poor robustness. In this paper, to address these issues, we propose a lightweight dense pedestrian detection model with finer-grained feature information interaction called MSCD-YOLO, which can achieve high accuracy, high performance and robustness with only a small number of parameters. In our model, the light-weight backbone network MobileViT is used to reduce the number of parameters while efficiently extracting both local and global features; the SCNeck neck network is designed to fuse the extracted features without losing information; and the DEHead detection head is utilized for multi-scale feature fusion to detect the targets. To demonstrate the effectiveness of our model, we conducted tests on the highly challenging dense pedestrian detection datasets Crowdhuman and Widerperson. Compared to the baseline model YOLOv8n, MSCD-YOLO achieved a 4.6% and 1.8% improvement in mAP@0.5, and a 5.3% and 2.6% improvement in mAP@0.5:0.95 on the Crowdhuman and Widerperson datasets, respectively. The experimental results show that under the same experimental conditions, MSCD-YOLO significantly outperforms the original model in terms of detection accuracy, efficiency, and model complexity.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001405280300001  
PubMed ID:  
39860808  
Indexed Date:  
2025-02-01  
Institution Address:  
Nanchang Hangkong Univ, Sch Informat Engn, Nanchang 330063, Peoples R China  
Institution:

### 646. Lightweight safety helmet detection algorithm using improved YOLOv5

摘要: In response to the challenges faced by existing safety helmet detection algorithms when applied to complex construction site scenarios, such as poor accuracy, large number of parameters, large amount of computation and large model size, this paper proposes a lightweight safety helmet detection algorithm based on YOLOv5, which achieves a balance between lightweight and accuracy. First, the algorithm integrates the Distribution Shifting Convolution (DSConv) layer and the Squeeze-and-Excitation (SE) attention mechanism, effectively replacing the original partial convolution and C3 modules, this integration significantly enhances the capabilities of feature extraction and representation learning. Second, multi-scale feature fusion is performed on the Ghost module using skip connections, replacing certain C3 module, to achieve lightweight and maintain accuracy. Finally, adjustments have been made to the Bottleneck Attention Mechanism (BAM) to suppress irrelevant information and enhance the extraction of features in rich regions. The experimental results show that improved model improves the mean average precision (mAP) by 1.0% compared to the original algorithm, reduces the number of parameters by 22.2%, decreases the computation by 20.9%, and the model size is reduced by 20.1%, which realizes the lightweight of the detection algorithm.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001262967900001  
Indexed Date:  
2024-07-12  
Institution Address:  
Tianjin Chengjian Univ, Sch Control & Mech Engn, Tianjin 300384, Peoples R China  
Institution Address:  
North China Univ Sci & Technol, Coll Elect Engn, Tangshan 063210, Hebei, Peoples R China  
Institution:  
Institution:

### 647. Bridging the Gap in Facial Age Progression: An Attention Mechanism

摘要: With the advent of Generative Adversarial Networks (GANs), significant progress has been made in facial aging prediction. However, existing methods still face considerable challenges. Many studies estimate the ages of individuals in images based on their birth dates rather than visual cues, leading to discrepancies between the predicted ages and the actual appearance of facial aging. Moreover, these approaches often overlook racial consistency, resulting in models predominantly tailored to European populations, which limits their generalizability across different races. To address these issues, we propose a novel facial aging prediction framework that employs three independent encoders to model identity, texture features, and facial skeletal structure. We replace traditional convolutional networks with an attention mechanism-based backbone, integrating spatial and channel attention mechanisms to capture both spatial relationships and age-related feature importance. These attention-enhanced feature maps are then processed through a pyramid feature fusion architecture to facilitate multi-scale feature extraction. Our model effectively captures the subtleties of facial aging across different demographics. Extensive experiments and ablation studies demonstrate that our approach excels in preserving identity, ensuring racial consistency, and generating realistic aging effects. The results further highlight the superior ability of the attention mechanisms to extract detailed, localized features essential for facial aging prediction.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001354547600001  
Indexed Date:  
2024-11-21  
Institution Address:  
Guangdong Ocean Univ, Fac Math & Comp Sci, Zhanjiang 524088, Peoples R China  
Institution:

### 648. DRMNet: A Multi-Task Detection Model Based on Image Processing for

摘要: To improve the efficiency and accuracy of autonomous driving vehicles' perception of the external environment, a multi-task detection model DRMNet (Dual-resolution Multi-task Network) is proposed that can be applied to autonomous driving scenarios, which can simultaneously complete the tasks of vehicle detection, lane detection, and drivable area detection. Firstly, given the loss of feature information by multiple downsampling in the backbone feature network, which affects the detection accuracy, the backbone of the model is designed as a two-pathway structure, which is used to extract shallow detail information and deep semantic information, respectively. Secondly, a multi-scale feature fusion module (MFFM) is designed to fuse the extracted shallow detail and deep semantic information. Then, different detection branches are designed according to the different characteristics of each detection task. Finally, Experiments on the BDD100K show the performance of DRMNet in three detection tasks: The recall and mAP of vehicle detection are 93.9% and 80.0% respectively. The accuracy of lane detection is 76.3%. The mIoU of drivable area detection is 92.2%. It is superior to the existing multi-task algorithm model, and the model has good generalization ability through actual scene experiments.  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
0018-9545  
eISSN:  
1939-9359  
入藏号:  
WOS:001132470500022  
Indexed Date:  
2024-03-12  
Institution Address:  
Beijing Jiaotong Univ, Sch Traff & Transportat, Beijing 100044, Peoples R China  
Institution Address:  
Beijing Jiaotong Univ, Minist Transport, Key Lab Transport Ind Big Data Applicat Technol C, Beijing 100044, Peoples R China  
Institution:  
Institution:

### 649. Multi-Scale Attention Feature Enhancement Network for Single Image

摘要: Aiming to solve the problem of color distortion and loss of detail information in most dehazing algorithms, an end-to-end image dehazing network based on multi-scale feature enhancement is proposed. Firstly, the feature extraction enhancement module is used to capture the detailed information of hazy images and expand the receptive field. Secondly, the channel attention mechanism and pixel attention mechanism of the feature fusion enhancement module are used to dynamically adjust the weights of different channels and pixels. Thirdly, the context enhancement module is used to enhance the context semantic information, suppress redundant information, and obtain the haze density image with higher detail. Finally, our method removes haze, preserves image color, and ensures image details. The proposed method achieved a PSNR score of 33.74, SSIM scores of 0.9843 and LPIPS distance of 0.0040 on the SOTS-outdoor dataset. Compared with representative dehazing methods, it demonstrates better dehazing performance and proves the advantages of the proposed method on synthetic hazy images. Combined with dehazing experiments on real hazy images, the results show that our method can effectively improve dehazing performance while preserving more image details and achieving color fidelity.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001082483200001  
PubMed ID:  
37836932  
Indexed Date:  
2023-10-28  
Institution Address:  
Changchun Univ Sci & Technol, Sch Optoelect Engn, Changchun 130022, Peoples R China  
Institution Address:  
Changchun Univ Sci & Technol, Zhongshan Inst, Zhongshan 528437, Peoples R China  
Institution:  
Institution:

### 650. T-Net: Deep Stacked Scale-Iteration Network for Image Dehazing

摘要: Haze reduces the visibility of image content and leads to failure in handling subsequent computer vision tasks. In this paper, we address the problem of single image dehazing by proposing a dehazing network named T-Net, which consists of a backbone network based on the U-Net architecture and a dual attention module. Multi-scale feature fusion can be achieved by using skip connections with a new fusion strategy. Furthermore, by repeatedly unfolding the plain T-Net, Stack T-Net is proposed to take advantage of the dependence of deep features across stages via a recursive strategy. To reduce network parameters, the intra-stage recursive computation of ResNet is adopted in our Stack T-Net. We take both the stage-wise result and the original hazy image as input to each T-Net and finally output the prediction of the clean image. Experimental results on both synthetic and real-world images demonstrate that our plain T-Net and the advanced Stack T-Net perform favorably against state-of-the-art dehazing algorithms and show that our Stack T-Net could further improve the dehazing effect, demonstrating the effectiveness of the recursive strategy.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
22  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
22  
ISSN:  
1520-9210  
eISSN:  
1941-0077  
入藏号:  
WOS:001102654000007  
Indexed Date:  
2023-12-16  
Institution Address:  
Shenzhen Univ, ATR Natl Key Lab Def Technol, Shenzhen 518060, Peoples R China  
Institution Address:  
Shenzhen Univ, Guangdong Key Lab Intelligent Informat Proc, Shenzhen 518060, Peoples R China  
Institution Address:  
Australian Natl Univ, Coll Engn & Comp Sci, Canberra, ACT 2601, Australia  
Institution Address:  
Sun Yat Sen Univ, Guangzhou 510275, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 651. Flame recognition and location algorithm based on YOLO v5s and binocular

摘要: An improved YOLO v5s framework in conjunction with binocular vision is developed to fulfill the stringent requirements for precise identification and accurate localization of flames. An Efficient Multi-scale Attention (EMA) mechanism is seamlessly integrated into the backbone network of the YOLO v5s model, thereby significantly improving the capacity to focus on and interpret critical target features. To bolster the multi-scale feature fusion capability, the Bidirectional Feature Pyramid Network (BiFPN) is strategically introduced at the fusion layer within the object detection architecture. The Semi-Global Block Matching (SGBM) algorithm is deployed to synchronize the binocular images, complemented by the least squares curve fitting method to adjust location inaccuracies, enhancing the precision of flame positioning. A binocular vision system is meticulously constructed to facilitate a series of rigorous experiments on flame detection and location. The experimental results demonstrate that the improved YOLO v5s network model achieves an exceptional flame recognition rate of 96.25%, which represents a noteworthy increase of 0.98% in accuracy in comparison to the original YOLO v5s model. Additionally, within a distance range of 3 to 10 meters, the flame ranging error is consistently less than 0.1 meters, successfully realizing the objective of high-precision flame detection and localization.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2631-8695  
入藏号:  
WOS:001424681400001  
Indexed Date:  
2025-02-27  
Institution Address:  
713 Res Inst CSSC, Zhengzhou 450015, Henan, Peoples R China  
Institution Address:  
Zhengzhou Key Lab Fire Protect Technol Special Pl, Zhengzhou 450015, Henan, Peoples R China  
Institution Address:  
Beihang Univ, Sch Mech Engn & Automat, Beijing 100191, Peoples R China  
Institution:  
713 Res Inst CSSC  
Institution:  
Zhengzhou Key Lab Fire Protect Technol Special Pl  
Institution:

### 652. A novel large-kernel residual grasp network for robot grasp detection

摘要: Efficient and accurate grasping remains a significant challenge due to the diverse object shapes, sizes, and poses encountered in practical scenarios. Previous grasp detection methods were limited regarding receptive fields. They have demonstrated insufficient capability in extracting relevant grasp features and have not effectively leveraged multi-scale features, resulting in limited detection accuracy. This paper introduces the large-kernel residual grasp network (LKRG-Net), a novel network designed to address these challenges by integrating advanced feature extraction and fusion techniques. Firstly, the proposed model utilizes a dual-encoding UniRepLKNet-ResNet50 backbone encoding grasp features at global and local levels, ensuring comprehensive extraction of relevant characteristics. Secondly, a grasp fusion splicing module effectively splices and merges the dual-encoded feature loss of crucial information. Finally, a selective fusion feature pyramid network decoding multi-scale feature information enhances the utilization of shallow features while altering the information. Comprehensive testing on the Cornell Jacquard and Jacquard\_V2 datasets shows that LKRG-Net surpasses existing advanced methods in accuracy and robustness. Grasping detection experiments conducted in real object scenarios further confirm the model's effectiveness in dynamic environments, providing a solid foundation for future advancements in robotic grasping tasks. The code can be obtained at the following address. https://github.com/Fyzyukk/LKRG-Net.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001385232900001  
Indexed Date:  
2025-01-05  
Institution Address:  
Harbin Univ Sci & Technol, Sch Automat, Heilongjiang Prov Key Lab Complex Intelligent Syst, Harbin 150080, Heilongjiang, Peoples R China  
Institution:

### 653. Mobilenetv2\_CA Lightweight Object Detection Network in Autonomous

摘要: A lightweight network target detection algorithm was proposed, based on MobileNetv2\_CA, focusing on the problem of high complexity, a large number of parameters, and the missed detection of small targets in the target detection network based on candidate regions and regression methods in autonomous driving scenarios. First, Mosaic image enhancement technology is used in the data pre-processing stage to enhance the feature extraction of small target scenes and complex scenes; second, the Coordinate Attention (CA) mechanism is embedded into the Mobilenetv2 backbone feature extraction network, combined with the PANet and Yolo detection heads for multi-scale feature fusion; finally, a Lightweight Object Detection Network is built. The experimental test results show that the designed network obtained the highest average detection accuracy of 81.43% on the Voc2007 + 2012 dataset, and obtained the highest average detection accuracy of 85.07% and a detection speed of 31.84 FPS on the KITTI dataset. The total amount of network parameters is only 39.5 M. This is beneficial to the engineering application of MobileNetv2 network in automatic driving.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
2227-7080  
入藏号:  
WOS:000979603000001  
Indexed Date:  
2023-05-16  
Institution Address:  
Anhui Polytech Univ, Sch Mech Engn, Wuhu 241000, Peoples R China  
Institution Address:  
Zhejiang Univ, Polytech Inst, Hangzhou 310027, Peoples R China  
Institution:  
Institution:

### 654. FashionSegNet: a model for high-precision semantic segmentation of

摘要: Clothing image segmentation is a method to predict the clothing category label of each pixel in the input image. We reduced the influence of the variability of image shots, the similarity of clothing categories, and the complexity of boundaries on the segmentation accuracy of clothing images by developing an advanced ResNet50-based semantic segmentation model in this study whose primary structure is the encoder-decoder. An improved spatial pyramid pooling module combined with a global feature extraction branch of a large convolution kernel is developed to achieve multi-scale feature fusion and improve the model's ability to identify clothing and its boundary features in different shots. Furthermore, to balance the clothing shape and category information in the model, a spatial and semantic information enhancement module is proposed, which can enhance the circulation of the information between different stages of the network through cross-stage connection technology. The model was finally trained and tested on the Deepfashion2 dataset. The comparison experiment demonstrates that the proposed model obtained the highest mIoU and Boundary IoU of 74.55% and 57.51%, respectively, compared with the DeepLabv3+, PSPNet, and other networks.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:000991071000002  
Indexed Date:  
2023-06-07  
Institution Address:  
Zhejiang Sci Tech Univ, Sch Mech Engn, 928-2 Main St,Xiasha Higher Educ Pk, Hangzhou 310018, Zhejiang, Peoples R China  
Institution:

### 655. Complex Text Detection Algorithm Based on Edge Attention Mechanism

摘要: In computer vision, automatically identifying and locating text in images or videos is an important task. Traditional text detection methods are not effective in detecting complex scenes in irregular rectangular areas such as text bending, spacing, and special shapes, which are mainly reflected in the fragmentation of text detection areas. In this paper, a text detection method based on edge attention mechanism is proposed to better adapt to complex scenes. The proposed method takes Encoder-Decoder as the core idea. First of all, an edge attention module is designed, including global attention and local attention. The global attention module is used to perceive the features of text regions and nontext regions, while the local attention module is used to learn the information of text boundaries. Then a multi-scale feature fusion process is designed, which can strengthen the edge information and key information of text regions. Finally, the model outputs probability maps and threshold maps, and generates high-precision binary maps of text regions. After experimental verification, the proposed method on the public data set significantly reduces the fragmentation of the detection area, improves the detection accuracy of the text area, and has better robustness for text detection scenes with unconventional rectangular areas.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
0218-1266  
eISSN:  
1793-6454  
入藏号:  
WOS:001154367900002  
Indexed Date:  
2024-02-15  
Institution Address:  
Chongqing Normal Univ, Coll Comp & Informat Sci, Chongqing 401331, Peoples R China  
Institution:

### 656. Dual-modal pedestrian detection method based on multi-scale feature

摘要: In the realms of urban safety,autonomous transportation,and intelligent surveillance,advanced pedestrian detection technology is crucial.Current methods based on visible light imaging face limitations in low-light or adverse weather conditions,leading to reduced detection accuracy.To address these challenges,this paper introduces a novel dual-modal pedestrian detection method,utilizing an enhanced YOLOv7 model supplemented by modal alignment(MA) and differential modal fusion(DMF) modules.These modules effectively harness dual-modal data,combining visible light and infrared imaging to improve detection performance under various environmental conditions.Experimental results indicate that the proposed method significantly enhances pedestrian detection accuracy across various scenarios,offering a promising solution for detection tasks in complex environments.  
摘要:  
在城市安全、自主交通和智能监控领域,先进的行人检测技术至关重要。当前基于可见光成像的方法在低光或恶劣天气条件下面临局限性,导致检测精度降低。为了应对这些挑战,该文提出了一种新型的双模态行人检测方法,该方法采用改进的YOLOv7模型,并增强了模态对齐(MA)和差分模态融合(DMF)模块。这些模块有效地利用双模态数据,结合可见光和红外成像,以改善各种环境条件下的检测性能。实验结果表明,所提出的方法在多种场景下均显著提高了行人检测的准确性,这为复杂环境下的检测任务提供了一种有潜力的解决方案。  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1005-9830  
入藏号:  
CSCD:7841166  
Indexed Date:  
2025-01-19  
Institution Address:  
南京理工大学自动化学院, 南京, 江苏 210094, 中国  
Institution Address:  
中国船舶集团有限公司第七○八研究所, 上海 200011, 中国  
Institution Address:  
School of Automation,Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China  
Institution Address:  
Marine Design and Research Institute of China, Shanghai 200011, China  
Institution:  
南京理工大学自动化学院  
Institution:  
中国船舶集团有限公司第七○八研究所  
Institution:  
School of Automation,Nanjing University of Science and Technology  
Institution:  
Marine Design and Research Institute of China

### 657. An Improved YOLOv5 Algorithm for Steel Surface Defect Detection

摘要: Scale of steel surface defects is different, but existing detection algorithms have poor multi-scale feature processing ability and low accuracy. Therefore, an improved YOLOv5 algorithm for steel surface defect detection is proposed. First, receptive field modules are added after the feature output layer of the backbone to enhance the discrimination and robustness of the features which can better perceive the feature information of different scales. Then, aligned feature aggregation modules are used to replace the traditional feature fusion structure to solve the feature misalignment problem in the fusion process of high and low resolution feature maps. Finally, decoupled heads with efficient channel attention mechanisms are used to output the detection results. The attention mechanism can adaptively calibrate the channel response, and the decoupled heads enable classification and regression tasks to be performed independently. The experimental results on NEU-DET dataset show that the mean average precision of the proposed method is 80.51%, which is 4.48% higher than that of the benchmark model, and the detection speed is 31.96 frame/s. Compared with other mainstream object detection algorithms, the proposed algorithm has higher accuracy while maintaining certain detection speed, enabling efficient steel surface defect detection.  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1006-4125  
入藏号:  
WOS:001180502700017  
Indexed Date:  
2024-03-20  
Institution Address:  
Tianjin Univ, Sch Microelect, Tianjin 300072, Peoples R China  
Institution Address:  
Tianjin Univ, Coll Intelligence & Comp, Tianjin 300072, Peoples R China  
Institution Address:  
Tianjin Key Lab Imaging & Sensing Microelect Tech, Tianjin 300072, Peoples R China  
Institution:  
Institution:  
Institution:  
Tianjin Key Lab Imaging & Sensing Microelect Tech

### 658. C3N: content-constrained convolutional network for mural image

摘要: Ancient murals, suffering from severe diseases, usually exhibit the absence or distortion of local areas. The damaged murals severely impaired people's visual appreciation and satisfaction in the digital conservation of cultural heritage. However, there is no large amount of murals due to their scarcity. In this paper, we propose a novel content-constrained convolutional network for mural image completion. This method employs frequency transformation to facilitate effective multi-scale feature fusion for image inpainting, taking into account both space and frequency domains. Our network uses adaptive space-varying activation functions to correct feature maps across scales. Our network also uses dual-domain partial convolution with a mask for computing on only valid points, whereas the mask is updated for the next layer. This iterative process is performed until the mask is filled to build the repaired image. The proposed method is verified on the datasets in comparison with baseline methods. The experimental results demonstrate that the proposed method achieves better results with less artifacts in repairing mural images and generally outperforms the state-of-the-art methods both quantitatively and qualitatively. The code and pretrained models are available at https://github.com/zhangyongqin/C3N.  
Web of Science 核心合集中的 "被引频次":  
5  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
0941-0643  
eISSN:  
1433-3058  
入藏号:  
WOS:000864595700001  
Indexed Date:  
2022-10-20  
Institution Address:  
Northwest Univ, Artificial Intelligence Res Inst, Xian 710127, Shaanxi, Peoples R China  
Institution Address:  
Northwest Univ, Art Sch, Xian 710127, Shaanxi, Peoples R China  
Institution Address:  
Northwest Univ, Sch Informat Sci & Technol, Xian 710127, Shaanxi, Peoples R China  
Institution Address:  
Shaanxi Hist Museum, Cultural Relics Protect Off, Xian, Shaanxi, Peoples R China  
Institution Address:  
Nottingham Trent Univ, Sch Sci & Technol, Nottingham NG11 8NS, England  
Institution:  
Institution:  
Institution:  
Institution:  
Shaanxi Hist Museum  
Institution:

### 659. Monitoring Anthropogenically Disturbed Parcels with Soil Erosion

摘要: Amidst burgeoning socioeconomic development, anthropogenic activities have exacerbated soil erosion. This erosion, characterized by its brief duration, high frequency, and considerable environmental degradation, presents a major challenge to ecological systems. Therefore, it is imperative to regulate and remediate erosion-prone, anthropogenically disturbed parcels, with dynamic change detection (CD) playing a crucial role in enhancing management efficiency. Currently, traditional methods for change detection, such as field surveys and visual interpretation, suffer from time inefficiencies, complexity, and high resource consumption. Meanwhile, despite advancements in remote sensing technology that have improved the temporal and spatial resolution of images, the complexity and heterogeneity of terrestrial cover types continue to limit large-scale dynamic monitoring of anthropogenically disturbed soil erosion parcels (ADPSE) using remote sensing techniques. To address this, we propose a novel ISegFormer model, which integrates the SegFormer network with a pseudo-residual multilayer perceptron (PR-MLP), cross-scale boundary constraint module (CSBC), and multiscale feature fusion module (MSFF). The PR-MLP module improves feature extraction by capturing spatial contextual information, while the CSBC module enhances boundary prediction through high- and low-level semantic guidance. The MSFF module fuses multiscale features with attention mechanisms, boosting segmentation precision for diverse change types. Model performance is evaluated using metrics, such as precision, recall, F1-score, intersection over union (IOU), and mean intersection over union (mIOU). The results demonstrate that our improved model performs exceptionally well in dynamic monitoring tasks for ADPSE. Compared to five other models, our model achieved an mIOU of 72.34% and a Macro-F1 score of 83.55% across twelve types of ADPSE changes, surpassing the other models by 1.52-2.48% in mIOU and 2.25-3.64% in Macro-F1 score. This work provides a theoretical and methodological foundation for policy-making in soil and water conservation departments.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001377622900001  
Indexed Date:  
2024-12-21  
Institution Address:  
China Univ Geosci Wuhan, Natl Engn Res Ctr Geog Informat Syst, Engn Res Ctr Nat Resource Informat Management & Di, Sch Geog & Informat Engn,Minist Educ, Wuhan 430074, Peoples R China  
Institution Address:  
China Univ Geosci, Sch Comp Sci, Wuhan 430074, Peoples R China  
Institution:  
Institution:

### 660. Multi-scale Lightweight Algorithm for UAV Aerial Target Detection

摘要: To address the high miss rates and low accuracy in detecting small objects in drone aerial images, caused by limited feature information and noise interference, we propose an improved algorithm, MLA-YOLO, based on YOLOv8s. First, a shallow feature enhancement network with a Global Context Block (GC-Block) is embedded in the backbone to mitigate noise interference and reduce small object feature loss during fusion. Second, deformable convolutions replace part of the standard convolutions in C2F to improve adaptability to geometric variations. Third, an ASPPF module, combined with average pooling, is introduced to enhance multi-scale feature representation and reduce miss rates. Finally, a middle-scale feature synthesis layer is embedded in the neck, with skip connections to ensure smoother transitions between feature scales and enhance feature reuse. Experiments on the VisDrone2019 and VOC2012 datasets show that MLA-YOLO achieves mAP@0.5 values of 40.1% and 72.5%, representing improvements of 8.4% and 3.2% over the baseline YOLOv8s. These results demonstrate the effectiveness and generalization of MLA-YOLO for small object detection in drone images.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1816-093X  
eISSN:  
1816-0948  
入藏号:  
WOS:001374721100012  
Indexed Date:  
2024-12-18  
Institution Address:  
Liaoning Inst Sci & Engn, Coll Informat Engn, Jinzhou 121010, Peoples R China  
Institution Address:  
Qinghai Normal Univ, Coll Comp, Xining 810008, Peoples R China  
Institution Address:  
Liaoning Tech Univ, Software Coll, Huludao 125105, Peoples R China  
Institution:  
Liaoning Inst Sci & Engn  
Institution:  
Institution:

### 661. A deep learning model based on improved feature pyramid networks for

摘要: Existing object detection models often use multi-scale feature fusion of feature pyramid to improve small object detection performance. However, in the shallow feature layer of the feature pyramid, the detection of small objects will be weakened due to the existence of large objects, and the semantic information of the upper feature layer will be lost due to lateral connection. To solve the above problems, an I-FPN feature pyramid is proposed. At the shallow feature layer, it erases the big object information and makes the model focus more on the small object. In the upper feature layer, the residual feature enhancement module is used to reduce the information loss. In addition, the model uses data augmentation techniques to improve the robustness. I-FPN feature pyramid was tested on VEDAI small target data set and PASCAL VOC universal target data set using Resnet master network. The experiment shows that, under the condition that the detection speed is not affected, the mAP index of VEDAI test set is increased by 2.4%, and that of VOC test set is increased by 0.5%.  
摘要:  
现有的目标检测模型常采用特征金子塔的多尺度特征融合来提升小目标检测性能。然而,在特征金字塔的浅层特征层,大目标的存在会削弱模型对小目标的检测,侧向连接会丢失高层特征层的语义信息。针对以上问题,提出了I-FPN特征金字塔。在浅层特征层,抹去大目标信息让模型更关注小目标;在高层特征层,使用残差特征增强模块减少信息损失。此外,模型还使用数据增广技术提升鲁棒性。 I-FPN特征金字塔使用Resnet为主干网络,在VEDAI小目标数据集和PASCAL VOC通用目标数据集上进行了实验。实验结果表明,在不影响检测速度的条件下,在VEDAI测试集上较原特征金字塔的mAP指标提升了2.4%,在VOC测试集上mAP指标提升了0.5%。  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1007-130X  
入藏号:  
CSCD:7462018  
Indexed Date:  
2023-08-12  
Institution Address:  
长沙理工大学计算机与通信工程学院, 长沙, 湖南 410114, 中国  
Institution Address:  
School of Computer &Communication Engineering,Changsha University of Science &Technology, Changsha, Hunan 410114, China  
Institution:  
长沙理工大学计算机与通信工程学院  
Institution:  
School of Computer &Communication Engineering,Changsha University of Science &Technology

### 662. Intelligent Blasthole Detection of Roadway Working Face Based on

摘要: Blasthole detection is crucial but challenging in tedious underground mining processes, given the diversity of surrounding rock backgrounds and uneven light intensity. However, existing algorithms have limitations in extracting image features and identifying differently sized objects. This study proposes a cascade-network-based blasthole detection method. The proposed method includes a blasthole feature extract transformation (BFET) module and a blasthole detection (BD) module. Firstly, we constructed the BFET module on the improved Cycle Generative Adversarial Network (CycleGAN) by multi-scale feature fusion. Then, we fused the convolution features of the generators in CycleGAN to obtain the enhanced feature map of the blasthole images. Secondly, the BD module was cascaded with the BFET module to accomplish the task of detecting blastholes. Results indicated that the detection accuracy of the blasthole image was significantly improved by strengthening the contrast of the image and suppressing over-exposure. The experimental results also showed that the proposed method enhanced the contrast of the image and could improve the accuracy of blasthole detection in real time. Compared with the YOLOv7 and CycleGAN+YOLOv7 methods, the detection accuracy of our method was improved by 5.34% and 2.38%, respectively.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
eISSN:  
2076-3417  
入藏号:  
WOS:001003464300001  
Indexed Date:  
2023-06-20  
Institution Address:  
China Univ Min & Technol Beijing, Sch Mech Elect & Informat Engn, Beijing 100083, Peoples R China  
Institution Address:  
China Univ Min & Technol Beijing, Sch Mech & Civil Engn, Beijing 100083, Peoples R China  
Institution:  
Institution:

### 663. Multi-Camera System: Imaging Enhancement and Application

摘要: A camera mimics the biological vision system for acquiring natural scenes. Human beings and other animals are mostly equipped with binocular stereo or multi-eye vision systems for high-fidelity scene acquisition, which motivates us to develop a multi-camera system to enhance the imaging capacity. Multi-dimensional and multi scale image/video acquisitions can be realized via heterogeneous sensors and shooting conditions of different cameras, and then leveraging matches across cameras to realize computational multiscale feature fusion for final enhanced reconstruction in respect of scale or dimensionality. In this article, we exemplify the multi-camera system in the applications of wide-field ultrahigh-definition imaging, high spatiotemporal video acquisition, high dynamicrange and low-light imaging enhancement to demonstrate its advantages in improving the imaging capacity, e. g., increasing imaging spatiotemporal resolution, expanding the field of view and extending imaging dynamic range. Key words imaging system; multicamera system; computational imaging; imaging enhancement; spatiotemporal resolution; dynamic range  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1006-4125  
入藏号:  
WOS:000752856200013  
Indexed Date:  
2022-02-25  
Institution Address:  
Nanjing Univ, Sch Elect Sci & Engn, Nanjing 210023, Jiangsu, Peoples R China  
Institution:

### 664. Multiscale feature fusion method for lane line detection based on time

摘要: In order to solve lane line detection in difficult traffic conditions, such as shadow occlusion, signpost degradation, curves, and tunnels, numerous models have been proposed. However, most of the existing models conduct independent single-frame image detection, which makes it difficult to utilize the continuity of driving images and is ineffective in challenging scenes. To this end, we suggest a spatiotemporal information processing model for lane line recognition that enhances critical features. In order to properly learn the correlation between continuous images, we first employ a convolutional gated recurrent unit to process spatiotemporal driving information on the basis of U-Net. Second, the pyramid split attention (PSA) module is used to enhance or suppress the obtained feature expressions. Finally, the skip connection is used to fuse the features of different scales encoded by each stage with the features processed by PSA and gradually restore to the original image size. Experiments on the TuSimple dataset demonstrate that our model outperforms representative lane line detection networks in challenging driving scenes, with an F-1-measure of up to 94.302%.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:000965365700057  
Indexed Date:  
2023-04-30  
Institution Address:  
Henan Univ Technol, Sch Artificial Intelligence & Big Data, Zhengzhou, Peoples R China  
Institution Address:  
Minist Educ, Key Lab Grain Informat Proc & Control, Zhengzhou, Peoples R China  
Institution Address:  
Henan Univ Technol, Sch Informat Sci & Engn, Zhengzhou, Peoples R China  
Institution:  
Institution:  
Minist Educ  
Institution:

### 665. AIE-YOLO: Auxiliary Information Enhanced YOLO for Small Object Detection

摘要: Small object detection is one of the key challenges in the current computer vision field due to the low amount of information carried and the information loss caused by feature extraction. You Only Look Once v5 (YOLOv5) adopts the Path Aggregation Network to alleviate the problem of information loss, but it cannot restore the information that has been lost. To this end, an auxiliary information-enhanced YOLO is proposed to improve the sensitivity and detection performance of YOLOv5 to small objects. Firstly, a context enhancement module containing a receptive field size of 21x21 is proposed, which captures the global and local information of the image by fusing multi-scale receptive fields, and introduces an attention branch to enhance the expressive ability of key features and suppress background noise. To further enhance the feature expression ability of small objects, we introduce the high- and low-frequency information decomposed by wavelet transform into PANet to participate in multi-scale feature fusion, so as to solve the problem that the features of small objects gradually disappear after multiple downsampling and pooling operations. Experiments on the challenging dataset Tsinghua-Tencent 100 K show that the mean average precision of the proposed model is 9.5% higher than that of the original YOLOv5 while maintaining the real-time speed, which is better than the mainstream object detection models.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
22  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
22  
eISSN:  
1424-8220  
入藏号:  
WOS:000882680000001  
PubMed ID:  
36365919  
Indexed Date:  
2022-11-24  
Institution Address:  
Xian Shiyou Univ, Sch Elect Engn, Xian 710065, Peoples R China  
Institution:

### 666. Cephalometric Mark Point Detection with Multi-scale Feature Fusion

摘要: The detection of cephalometric mark point is extremely crucial for clinical diagnosis,treatment planning,and research.To improve the accuracy of automatic detection,an improved multi-scale feature fusion detection model,AIW-Net, is proposed.AIW-Net uses the pre-trained lightweight network MobileNetV2 as the backbone network for feature extraction. The intermediate module comprises up and down sampling paths,and the improved inverse residual module is used to reduce the feature loss in the down-sampling process.The intermediate supervision from coarse to fine is introduced into the decoder module,the obtained multi-scale heat map is fused with the feature map,and the attention gate is used in the jump connection to effectively suppress the response of the background region in the feature map.The experimental results of the benchmark dataset Test 1 provided by ISBI 2015 Grand Challenge show that the proposed model achieves a Mean Radial Error(MRE) of 1.14 mm,and a Successful Detection Rate(SDR) of 86.38% and 92.10% within the clinically acceptable error range of 2 mm and 2.5 mm,respectively,which are better than that obtained by W-Net,IW-Net and other models.  
摘要:  
头影标志点检测对于临床诊断、治疗计划和研究至关重要。为提高自动检测的准确性,提出一种改进的多尺度特征融合检测模型AIW-Net。采用经过预训练的轻量型网络MobileNetV2作为主干网络进行特征提取,使用上、下采样路径构成中间模块,利用改进的倒残差模块减少下采样过程中的特征损失。在解码器模块中引入从粗到细的中间监督,将得到的多个尺度热图与特征图进行融合,并在跳跃连接中使用注意力门,有效抑制特征图中的背景区域响应。在ISBI 2015 Grand Challenge提供的基准数据集Test 1上进行实验,结果表明,该模型的平均径向误差为1.14 mm,在临床可接受的误差范围2 mm与2.5 mm内的成功检测率分别为86.38%与92.10%,性能优于W-Net、IW-Net等模型。  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1000-3428  
入藏号:  
CSCD:7429552  
Indexed Date:  
2023-07-01  
Institution Address:  
中北大学大数据学院, 太原, 山西 030051, 中国  
Institution Address:  
太原学院智能与自动化系, 太原, 山西 030032, 中国  
Institution Address:  
山西医科大学口腔医学院, 太原, 山西 030001, 中国  
Institution Address:  
College of Big Data,North University of China, Taiyuan, Shanxi 030051, China  
Institution Address:  
Department of Intelligence and Automation, Taiyuan University, Taiyuan, Shanxi 030032, China  
Institution Address:  
School of Stomatology,Shanxi Medical University, Taiyuan, Shanxi 030001, China  
Institution:  
中北大学大数据学院  
Institution:  
太原学院智能与自动化系  
Institution:  
山西医科大学口腔医学院  
Institution:  
College of Big Data,North University of China  
Institution:  
Department of Intelligence and Automation, Taiyuan University  
Institution:  
School of Stomatology,Shanxi Medical University

### 667. Parallel Multi-Scale Feature Fusion Algorithm for Thermal Tomography

摘要: The application of thermal tomography in the detection of medical breast tumors has an important application prospect,but in the process of clinical application,doctors are prone to misdiagnosis through thermal tomography images.Therefore,an intelligent segmentation algorithm is proposed to assist diagnosis.However,due to the lack of data and small proportion of lesion area,medical thermal tomography images rely on classical segmentation models,such as FCN and U-Net,which are prone to problems such as discontinuous segmentation and unrefined segmentation of boundary details.Therefore,a semantic segmentation model based on parallel multi-scale feature fusion is designed.Through repeated information exchange among parallel multi-resolution feature subnets,the model can not only ensure the accuracy of semantic information of segmentation results,but also effectively capture the detail features of the focal area.The proposed method achieves the mean intersection over union of 0.6357 on the dataset of thermal tomography medical images,an improvement of 5.14 percentage points compared with the classic U-NET segmentation network,and has a better performance in the segmentation of small target areas such as lumps and blood vessels.The experimental results show that the proposed algorithm has practical significance for the clinical application of thermal tomography in the auxiliary diagnosis of breast cancer.  
摘要:  
热层析技术应用于医学乳腺肿瘤检测具有重要的应用前景,但是医生在临床应用过程中,通过热层析图像诊断容易出现主观差异性误诊现象,为此,提出了一种智能化分割算法用于辅助诊断。然而,医学热层析图像由于目前数据量匮乏,且病灶区域占比小,依靠经典的分割模型,如FCN、U-Net容易出现分割不连续,边界细节分割不精细等问题。设计了一种基于并行多尺度特征融合的语义分割模型,模型通过并行的多分辨率特征子网之间反复的信息交换,在保证分割结果语义信息准确之外,还能有效地抓取病灶区域的细节特征。该方法在热层析医学图像数据集上取得了0.6357的均交并比,相较于经典的U-Net分割网络,取得了5.14个百分点的提升,在肿块和血管等小目标区域的细节分割上有着更出色的表现。实验结果表明,该算法对热层析临床用于乳腺癌的辅助诊断具有现实意义。  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7314708  
Indexed Date:  
2023-02-04  
Institution Address:  
武汉大学物理科学与技术学院, 武汉, 湖北 430072, 中国  
Institution Address:  
School of Physics Science and Technology,Wuhan University, Wuhan, Hubei 430072, China  
Institution:  
武汉大学物理科学与技术学院  
Institution:  
School of Physics Science and Technology,Wuhan University

### 668. Multi-Scale Feature Refined Network for Human Pose Estimation

摘要: Occlusive keypoints has been a challenge for human pose estimation, especially the mutual occlusion of human bodies. One possible solution to this problem is to utilize multi-scale features, where small scale features are capable of identifying keypoints, while large-scale features can capture the relationship between keypoints. Feature fusion among multi-scale features allows for the exchange of information between keypoints, facilitating the inference of occluded keypoints based on the identified keypoints. However, it's found that there are invalid features in feature fusion which will interfere valid feature. In this paper, we propose multi-scale feature refined network (MSFRNet) based on HRNet and a new attention module namely multi-resolution attention module (MRAM). The proposed MRAM is designed to strengthen the effective information while suppressing redundant information. It has multiple inputs and outputs and can learn the relationships between keypoints while retaining detailed information. The proposed MSFRNet outperforms HRNet, achieving a 1.4AP improvement on the COCO dataset with only a marginal computational increase of 0.35 GFLOPs. Additionally, it demonstrates superior performance with a 0.9AP, 0.7AP, and 1.8AP improvement on the MPII, CrowdPose and OCHuman datasets, respectively. Furthermore, compared with the latest attention mechanism PSA, the MSFRNet exhibits lower computational cost while maintaining the same pose-estimation accuracy.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0218-0014  
eISSN:  
1793-6381  
入藏号:  
WOS:001106865300001  
Indexed Date:  
2023-12-14  
Institution Address:  
Beijing Univ Chem Technol, Coll Informat Sci & Technol, Beijing, Peoples R China  
Institution:

### 669. Dual-Path and Multi-Scale Enhanced Attention Network for Retinal

摘要: Early computer-aided early diagnosis (CAD) based on retinal imaging is critical to the timely management and treatment planning of retina-related diseases. However, the inherent characteristics of retinal images and the complexity of their pathological patterns, such as low image contrast and different lesion sizes, restrict the performance of CAD systems. Recently, ultra-wide-field (UWF) retinal images have become a useful tool for disease detection due to the capability of capturing much broader view of retina (i.e., up to 200 degrees), in comparison with the most commonly used retinal fundus images (45 degrees). In this paper, we propose an attention-based multi-branch network for the diseases classification of four different subject groups. The proposed method consists of a multi-scale feature fusion module and a dual attention module. Specifically, small-scale lesions are identified using the features extracted from the multi-scale feature fusion module. To better explore the obtained features, the dual attention module with a global attention graph is incorporated to enable the network to recognize the salient objects of interest. Comprehensive validations on both private and public datasets were carried out to verify the effectiveness of the proposed model.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
ISSN:  
2169-3536  
入藏号:  
WOS:000988213400001  
Indexed Date:  
2023-05-26  
Institution Address:  
Zhejiang Univ Technol, Sch Mech & Engn, Hangzhou 310014, Peoples R China  
Institution Address:  
Chinese Acad Sci, Cixi Inst Biomed Engn, Ningbo Inst Mat Technol & Engn, Ningbo 315399, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Sch Biomed Engn, Shenzhen 510275, Peoples R China  
Institution Address:  
Wenzhou Med Univ, Affiliated Ningbo Eye Hosp, Ningbo 325035, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 670. Exploring multi-scale forgery clues for stereo super-resolution image

摘要: Existing forgery image localization methods have achieved impressive performance on monocular images struggle to maintain comparable performance on stereo super-resolution (SR) images. The stereo image process usually introduces additional noise, i.e., reconstruction artifacts, which interferes with the extraction of forgery clues and thus undermines the localization performance. To mitigate this issue, we present a multi scale attention framework named SSR-IFL for stereo SR image forgery localization, which explores multi-scale forgery clues in stereo SR images to minimize disturbances caused by reconstruction artifacts. Specifically, novel framework integrates two significant components: the multi-scale feature extraction (MSFE) network thoroughly mine forgery clues against the inference of reconstruction artifacts, and the bidirectional progressive feature fusion (BPFF) network to effectively reason the learned forgery clues and obtain robust forgery feature representation. Under the MSFE network, the multi-dilated self-attention (MDSA) block makes full use of multi resolution pixel correlation to adapt to changes in pixel correlation distributions induced by the stereo reconstruction process, thus extracting more precise multi-scale forgery clues. The BPFF network introduces bidirectional feature fusion strategy to fuse multi-scale forgery clues, effectively eliminating redundant noise such as reconstruction artifacts and enhancing the robustness of forgery feature representation. Experimental results on public image datasets demonstrate that our method can achieve not only competitive performance on stereo SR images but also on monocular images.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:001373588400001  
Indexed Date:  
2024-12-16  
Institution Address:  
Sun Yat Sen Univ, Sch Comp Sci & Engn, MoE Key Lab Informat Technol, Guangdong Prov Key Lab Informat Secur Technol, Guangzhou 510006, Peoples R China  
Institution:

### 671. SDNet: Image-based sonar detection network for multi-scale objects

摘要: Autonomous Underwater Vehicle (AUV) carrying sonar for object detection has become one of the main ways of ocean exploration. However, object detection in sonar images always faces the problems of balancing detection accuracy and efficiency. To this end, this paper proposes an efficient underwater object detection network for sonar images named SDNet. In the model, the authors construct a new feature extraction network based on RepVGG to balance the detection accuracy and speed. By combining channel attention with RepVGG, useful information of high-order feature maps captured can be selectively paid attention to. Then, the authors design a feature fusion network to efficiently converge the location and semantic information of multi-scale feature maps. In the network, the authors propose a lightweight cross stage partial network for sonar (CSP\_S) module suitable for sonar images, which can enhance the model's feature fusion capability and simplify the model.Finally, to reduce the conflict between classification and regression tasks, the authors leverage the Decoupled Head for the sonar object classification and localization. By testing on the self-built Underwater Sonar Dataset underwater sonar dataset (USD) and the public sonar dataset sonar common target detection dataset (SCTD), the detection accuracy of SDNet reaches 99.52% and 95.20%, respectively. Moreover, the detection speed reaches 114 frames per second (FPS) and 138 FPS, respectively. The experimental results show that SDNet can effectively balance the sonar detection accuracy and efficiency.Code is available at .  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1751-9659  
eISSN:  
1751-9667  
入藏号:  
WOS:000899450100001  
Indexed Date:  
2023-01-03  
Institution Address:  
Hohai Univ, Coll Internet Things Engn, Changzhou, Jiangsu, Peoples R China  
Institution Address:  
Hohai Univ, Coll Comp & Informat, Nanjing, Jiangsu, Peoples R China  
Institution Address:  
Hohai Univ, Coll Internet Things Engn, Changzhou 213022, Jiangsu, Peoples R China  
Institution:  
Institution:  
Institution:

### 672. UAV detection algorithm based on spatial correlation enhancement

摘要: A small target detection method for unmanned aerial vehicle (UAV) based on adaptive up-sampling and spatial correlation enhancement was proposed,to resolve the problem of false detection and missed detection caused by the small size of UAV and the difficulty of feature extraction under complex backgrounds.Firstly,the important contextual information was obtained by multi-scale dilated convolution,and then the attention feature fusion module was used to suppress the information conflict of multi-scale feature fusion;Secondly,a new up-sampling method of sub-pixel convolution and bilinear interpolation adaptive fusion was adopted to balance the computation and to fuse more UAV feature information;Finally,spatial correlation enhancement strategies for local and global spatial features were performed on deep features to improve the sensitivity of foreground targets in complex backgrounds and enhance target expression to suppress background noise.Ablation experiments and comparative experiments were implemented on the self-made UAV dataset.The mAP0.5 and mAP0.5:0.95 of the proposed algorithm were increased by 2.4% and 2.7% respectively,compared with those of the original YOLOv5 algorithm.Furthermore,the detection speed was able to achieve 58.5 frames per second.The performance of the proposed algorithm was also verified on the VisDrone2019 dataset,and its mAP0.5 and mAP0.5:0.95 were respectively higher than those of the YOLOv5 algorithm by 4.6% and 1.3%.  
摘要:  
针对无人机(UAV)体积小、复杂背景下特征难以提取导致被误检和漏检的问题,提出基于自适应上采样和空间相关性增强的无人机小目标检测方法.采用多尺度的空洞卷积获取重要的上下文信息,然后通过注意力特征融合模块抑制多尺度特征融合造成的信息冲突;采用亚像素卷积和双线性插值自适应融合的新上采样方式,融合更多无人机特征信息,同时平衡计算量;对深层特征图的空间局部特征和全局特征采用空间相关性增强策略,提高复杂背景下前景目标的敏感度,增强目标表达和抑制背景噪声.在自制无人机数据集上进行消融实验和对比实验,与原始YOLOv5算法相比,本算法的mAP0.5和mAP0.5∶0.95分别提高了2.4%和2.7%,检测速度能够达到58.5帧/s;在VisDrone2019数据集上进行验证,本算法较YOLOv5算法的mAP0.5和mAP0.5∶0.95分别提高了4.6%和1.3%.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1008-973X  
入藏号:  
CSCD:7669366  
Indexed Date:  
2024-05-17  
Institution Address:  
河南工业大学电气工程学院;;北京理工大学自动化学院, ;;, 郑州;;, 河南;;北京 450001;;100081, 中国  
Institution Address:  
河南工业大学电气工程学院, 郑州, 河南 450001, 中国  
Institution Address:  
College of Electrical Engineering,Henan University of Technology;;School of Automation,Beijing Institute of Technology, ;;, Zhengzhou;;, ;;Beijing 450001;;100081  
Institution Address:  
College of Electrical Engineering,Henan University of Technology, Zhengzhou, Henan 450001, China  
Institution:  
河南工业大学电气工程学院;;北京理工大学自动化学院  
Institution:  
河南工业大学电气工程学院  
Institution:  
College of Electrical Engineering,Henan University of Technology;;School of Automation,Beijing Institute of Technology  
Institution:  
College of Electrical Engineering,Henan University of Technology

### 673. Mobile monocular depth estimation based on multi-scale feature fusion

摘要: The current depth estimation model based on depth learning has a large number of parameters, which is difficult to adapt to mobile devices. To address this issue, a lightweight depth estimation method with multi-scale feature fusion that can be deployed on mobile devices is proposed. Firstly, MobileNetV2 is used as the backbone to extract features of four scales. Then, by constructing skip connection paths from the encoder to the decoder, the features of the four scales are fused, fully utilizing the combined positional information from lower layers and semantic information from higher layers. Finally, the fused features are processed through convolutional layers to produce high-precision depth images. After training and testing on NYU Depth Dataset V2, the experimental results show that the proposed model achieves advanced performance with an evaluation index of delta1 up to 0.812 while only having 1.6\*10~6 parameters numbers. Additionally, it only takes 0.094 seconds to infer a single image on the Kirin 980 CPU of a mobile device, demonstrating its practical application value.  
摘要:  
目前基于深度学习的深度估计模型参数量大,难以适应移动端设备。针对此问题,提出一种可以部署在移动端的多尺度特征融合轻量级深度估计方法。首先,以MobileNetV2为主干,提取出4个尺度的特征。然后,通过构建编码器到解码器的跳跃连接路径,将4个尺度的特征进行融合,充分利用融合低层的位置信息和高层的语义信息。最后,融合后的特征通过卷积层得出高精度的深度图像。在NYU Depth Dataset V2数据集上进行了训练和测试,结果表明,该模型的参数量在仅有1.6\*10~6的情况下,评估指标delta1高达0.812,在移动端的麒麟980 CPU上推理一幅图像仅需要0.094 s,具有实际应用价值。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1007-130X  
入藏号:  
CSCD:7825901  
Indexed Date:  
2025-01-07  
Institution Address:  
广西大学计算机与电子信息学院, 南宁, 广西 530004, 中国  
Institution Address:  
广西大学计算机与电子信息学院;;广西多媒体通信与网络技术重点实验室, ;;广西多媒体通信与网络技术重点实验室, 南宁;;南宁, 广西;;广西 530004;;530004, 中国  
Institution Address:  
School of Computer and Electronics Information,Guangxi University, Nanning, Guangxi 530004, China  
Institution Address:  
School of Computer and Electronics Information,Guangxi University;;Guangxi Key Laboratory of Multimedia Communications and Network Technology, ;;Guangxi Key Laboratory of Multimedia Communications and Network Technology, Nanning;;Nanning, ;; 530004;;530004  
Institution:  
广西大学计算机与电子信息学院  
Institution:  
广西大学计算机与电子信息学院;;广西多媒体通信与网络技术重点实验室  
Institution:  
School of Computer and Electronics Information,Guangxi University  
Institution:  
School of Computer and Electronics Information,Guangxi University;;Guangxi Key Laboratory of Multimedia Communications and Network Technology

### 674. Obstacle detection method for guide system based on CE-YOLOX

摘要: Travel has always been a problem for blind people due to their lack of visual perception of the surrounding environment. This paper presents an improved obstacle detection algorithm CE-YOLOX based on YOLOX for guide system. Firstly, in order to reduce semantic information loss caused by feature fusion network when feature channel is reduced, sub-pixel hopping fusion module SSF and sub-pixel context enhancement module SCE are used to make full use of channel information and semantic information of different scales, and channel attention guide module CAG is used to reduce aliasing effect caused by multi-scale feature fusion. Secondly, in order to make the model more focused on effective features, the global attention mechanism GAM is introduced to improve the performance of the model by reducing the information dispersion and amplifying the global interactive representation. Then, the position regression function IOU-LOSS of the original model is replaced by SIOU-LOSS, which speeds up the regression speed and precision of the frame. Finally, the detection platform of the guide system is built and the proposed algorithm is transplanted to the edge computing device NVIDIA Xavier NX. The experimental results show that the obstacle algorithm of the improved guide system has the same mAP on the server and NVIDIA Xavier NX platform, which is improved to 90.53%, 2.45% higher than the original YOLOX model algorithm. The detection speed reaches 75.93 FPS on the server. The model in this paper not only gives consideration to the detection speed but also improves the accuracy, which is significantly better than the comparison algorithm. It meets the requirements of edge computing equipment and has practical application value.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1007-2780  
eISSN:  
2097-3217  
入藏号:  
WOS:001143657800014  
Indexed Date:  
2024-02-03  
Institution Address:  
Guizhou Univ, Coll Big Data & Informat Engn, Guiyang 550025, Peoples R China  
Institution:

### 675. Improved Model for Smoke Detection Based on Concentration Features using

摘要: Smoke is often present in the early stages of a fire. Detecting low smoke concentration and small targets during these early stages can be challenging. This paper proposes an improved smoke detection algorithm that leverages the characteristics of smoke concentration using YOLOv7tiny. The improved algorithm consists of the following components: 1) utilizing the dark channel prior theory to extract smoke concentration characteristics and using the synthesized alpha RGB image as an input feature to enhance the features of sparse smoke; 2) designing a light-BiFPN multi-scale feature fusion structure to improve the detection performance of small target smoke; 3) using depth separable convolution to replace the original standard convolution and reduce the model parameter quantity. Experimental results on a self-made dataset show that the improved algorithm performs better in detecting sparse smoke and small target smoke, with mAP@0.5 and Recall reaching 94.03% and 95.62% respectively, and the detection FPS increasing to 118.78 frames/s. Moreover, the model parameter quantity decreases to 4.97M. The improved algorithm demonstrates superior performance in the detection of sparse and small smoke in the early stages of a fire.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2158-107X  
eISSN:  
2156-5570  
入藏号:  
WOS:001084247800001  
Indexed Date:  
2023-11-02  
Institution Address:  
Zhengzhou Univ Light Ind, Zhengzhou 450000, Henan, Peoples R China  
Institution Address:  
SIPPR Engn Grp Co Ltd, Zhengzhou 450007, Henan, Peoples R China  
Institution Address:  
Zhengzhou Univ Ind Technol, Zhengzhou 451100, Henan, Peoples R China  
Institution:  
Institution:  
SIPPR Engn Grp Co Ltd  
Institution:  
Zhengzhou Univ Ind Technol

### 676. NLNet: A narrow-channel lightweight network for finger multimodal

摘要: Multimodal biometric recognition has attracted more and more attention in recent years because of its security and accuracy. Compared with the single use of fingerprint or finger vein feature recognition, the multi-modal feature recognition method based on fingerprint and finger vein significantly improves the recognition performance. However, most of the multi-modal feature recognition networks have the disadvantages of large number of parameters and high training cost. In this paper, a narrow-channel lightweight network NLNet for fingerprint and finger vein recognition is proposed. The network adopts asymmetric narrow channel structure for lightweight design, and combines shallow network to improve the discriminating nature of the extracted features, which significantly reduces the model parameters and computation. In addition, a lightweight feature extraction module for building feature extraction branches is designed for NLNet. This module takes dimensional transformation feature extraction as the backbone, and the joint extension module and attention mechanism obtain low-redundancy multi-scale feature information. In terms of feature fusion, a feature fusion method based on PatchPooling is proposed. This method combines the characteristics of modal images, and uses Spatial dimension local mapping to increase the utilization rate of low-dimensional features, which effectively improves the richness of classified features. In this paper, experiments were carried out on the SDUMLA-HMT, NUPT-FPV, FVC-HKP and HDPR-310 multimodal finger datasets, and the recognition accuracy was high as 97.72 %, 99.10 %, 99.67 % and 99.74 %, respectively. In addition, the effectiveness of the model is verified by comparing with other advanced methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
1051-2004  
eISSN:  
1095-4333  
入藏号:  
WOS:001232726200001  
Indexed Date:  
2024-06-03  
Institution Address:  
Heilongjiang Univ, Coll Elect Engn, Harbin 150080, Peoples R China  
Institution:

### 677. A Model for Infrastructure Detection along Highways Based on Remote

摘要: Infrastructure along the highway refers to various facilities and equipment: bridges, culverts, traffic signs, guardrails, etc. New technologies such as artificial intelligence, big data, and the Internet of Things are driving the digital transformation of highway infrastructure towards the future goal of intelligent roads. Drones have emerged as a promising application area of intelligent technology in this field. They can help achieve fast and precise detection, classification, and localization of infrastructure along highways, which can significantly enhance efficiency and ease the burden on road management staff. As the infrastructure along the road is exposed to the outdoors for a long time, it is easily damaged and obscured by objects such as sand and rocks; on the other hand, based on the high resolution of the images taken by Unmanned Aerial Vehicles (UAVs), the variable shooting angles, complex backgrounds, and high percentage of small targets mean the direct use of existing target detection models cannot meet the requirements of practical applications in industry. In addition, there is a lack of large and comprehensive image datasets of infrastructure along highways from UAVs. Based on this, a multi-classification infrastructure detection model combining multi-scale feature fusion and an attention mechanism is proposed. In this paper, the backbone network of the CenterNet model is replaced with ResNet50, and the improved feature fusion part enables the model to generate fine-grained features to improve the detection of small targets; furthermore, the attention mechanism is added to make the network focus more on valuable regions with higher attention weights. As there is no publicly available dataset of infrastructure along highways captured by UAVs, we filter and manually annotate the laboratory-captured highway dataset to generate a highway infrastructure dataset. The experimental results show that the model has a mean Average Precision (mAP) of 86.7%, an improvement of 3.1 percentage points over the baseline model, and the new model performs significantly better than other detection models overall.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
eISSN:  
1424-8220  
入藏号:  
WOS:000977576200001  
PubMed ID:  
37112188  
Indexed Date:  
2023-05-14  
Institution Address:  
Xinjiang Univ, Sch Informat Sci & Engn, Urumqi 830017, Peoples R China  
Institution:

### 678. YOLOv5-TS: Detecting traffic signs in real-time

摘要: Traffic sign detection plays a vital role in assisted driving and automatic driving. YOLOv5, as a one-stage object detection solution, is very suitable for Traffic sign detection. However, it suffers from the problem of false detection and missed detection of small objects. To address this issue, we have made improvements to YOLOv5 and subsequently introduced YOLOv5-TS in this work. In YOLOv5-TS, a spatial pyramid with depth-wise convolution is proposed by replacing maximum pooling operations in spatial pyramid pooling with depth-wise convolutions. It is applied to the backbone to extract multi-scale features at the same time prevent feature loss. A Multiple Feature Fusion module is proposed to fuse multi-scale feature maps multiple times with the purpose of enhancing both the semantic expression ability and the detail expression ability of feature maps. To improve the accuracy in detecting small even extra small objects, a specialized detection layer is introduced by utilizing the highest-resolution feature map. Besides, a new method based on k-means++ is proposed to generate stable anchor boxes. The experiments on the data set verify the usefulness and effectiveness of our work.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
2296-424X  
入藏号:  
WOS:001116312800001  
Indexed Date:  
2024-01-12  
Institution Address:  
Henan Polytech Univ, Sch Software, Jiaozuo, Peoples R China  
Institution Address:  
Anyang Inst Technol, Anyang, Peoples R China  
Institution Address:  
Henan Polytech Univ, Sch Comp Sci & Technol, Jiaozuo, Peoples R China  
Institution:  
Institution:  
Institution:

### 679. Segmentation of Heart Sound Signal Based on Multi-Scale Feature Fusion

摘要: Analyzing heart sound signals presents a novel approach for early diagnosis of pediatric congenital heart disease. The existing segmentation algorithms have limitations in accurately distinguishing the first (S1) and second (S2) heart sounds, limiting the diagnostic utility of cardiac cycle data for pediatric pathology assessment. This study proposes a time bidirectional long short-term memory network (TBLSTM) based on multi-scale analysis to segment pediatric heart sound signals according to different cardiac cycles. Mel frequency cepstral coefficients and dynamic characteristics of the heart sound fragments were extracted and input into random forest for multi-classification of congenital heart disease. The segmentation model achieved an overall F1 score of 94.15% on the verification set, with specific F1 scores of 90.25% for S1 and 86.04% for S2. In a situation where the number of cardiac cycles in the heart sound fragments was set to six, the results for multi-classification achieved stabilization. The performance metrics for this configuration were as follows: accuracy of 94.43%, sensitivity of 95.58%, and an F1 score of 94.51%. Furthermore, the segmentation model demonstrates robustness in accurately segmenting pediatric heart sound signals across different heart rates and in the presence of noise. Notably, the number of cardiac cycles in heart sound fragments directly impacts the multi-classification of these heart sound signals.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2306-5354  
入藏号:  
WOS:001322902700001  
PubMed ID:  
39329618  
Indexed Date:  
2024-10-07  
Institution Address:  
Jiangsu Univ, Res Ctr Fluid Machinery Engn & Technol, Zhenjiang 212013, Peoples R China  
Institution Address:  
Texas Tech Univ, Dept Mech Engn, Lubbock, TX 79411 USA  
Institution:  
Institution:

### 680. Remote Sensing Rotating Object Detection Based on Multi-Scale Feature

摘要: A rotation remote sensing target detection algorithm based on multi -scale feature extraction is proposed, because high -resolution remote sensing images have large object scale differences, dense small -object arrangements, and strong orientation. In this study, CenterNet was chosen as the benchmark model and redesigned. First, to improve the context information extraction ability, we proposed and applied the receptive field expansion module combined with multi -scale cavity convolution. Second, the extraction ability of the algorithm for multi -scale targets was improved in combination with adaptive feature fusion. Finally, we redesigned the CenterNet detection head and updated the loss function to improve the detection performance of the model for rotating objects. The designed model is named CenterNet for remote sensing images (CenterNet-RS). Experiments were performed on the DOTA dataset, and the mean average precision (mAP) of CenterNet-RS reaches 73. 01%, which is 9. 45 percentage points higher than the baseline model. Thus, the experimental findings demonstrate that the proposed method can significantly increase the target detection accuracy for remote sensing images.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
被引频次合计:  
6  
ISSN:  
1006-4125  
入藏号:  
WOS:001020949300010  
Indexed Date:  
2023-08-12  
Institution Address:  
Beijing Informat Sci & Technol Univ, Key Lab Modern Measurement & Control Technol, Minist Educ, Beijing 100089, Peoples R China  
Institution Address:  
Beijing Informat Sci & Technol Univ, Mech Elect Engn Sch, Beijing 100089, Peoples R China  
Institution:  
Institution:

### 681. Improved YOLOv3 Helmet Wearing Detection Method

摘要: Aiming at the problem of low accuracy and slow detection rate of helmet wearing detection in intelligent monitoring,a helmet wearing detection algorithm YOLOv3-WH based on improved YOLOv3(You Only Look Once) is proposed.The network structure on the basis of the YOLOv3 algorithm is improved,the scale of the input image is increased.Using deep separable convolution structure can replace the Darknet-53 traditional convolution,reduce the loss of features,model parameters,and increase the detection rate.Using multi-scale feature detection can increase the shallow detection scale,add 4 times the up-sampling feature fusion structure and improve the accuracy of helmet wearing detection.Optimizing the K-Means clustering algorithm can obtain the anchor box of the helmet wearing detection,according to the predicted scale size allocate suitable anchors,the model training and detection rate are improved.Experimental results show that compared with YOLOv3,YOLOv3-WH has a 64% increase in frame detection per second(FPS),and an average detection accuracy(mAP) has increased by 6.5%.The algorithm improves the detection rate of helmet wearing while improving detection.The accuracy rate is of certain practicality for the detection of helmet wearing.  
摘要:  
针对在智能监控中安全帽佩戴检测准确率低和检测速率慢的问题,提出一种基于改进YOLOv3(You Only Look Once)的安全帽佩戴检测算法YOLOv3-WH。在YOLOv3算法的基础上改进网络结构,增大输入图像的尺度,使用深度可分离卷积结构替换Darknet-53传统卷积,减少特征的丢失,缩减模型参数,提升检测速率;使用多尺度特征检测,增加浅层检测尺度,添加4倍上采样特征融合结构,提高安全帽佩戴检测准确率;优化K-Means聚类算法,获取安全帽佩戴检测的先验框(anchor box),按照预测尺度大小分配适合的anchor,提升模型训练和检测速率。实验结果表明YOLOv3-WH相比YOLOv3,每秒检测帧数(FPS)提高了64%,检测平均精确度(mAP)提高了6.5%,该算法在提升了安全帽佩戴检测速率的同时提升了检测的准确率,对安全帽佩戴检测具有一定的实用性。  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
8  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
1002-8331  
入藏号:  
CSCD:6992885  
Indexed Date:  
2021-09-10  
Institution Address:  
四川轻化工大学自动化与信息工程学院, 自贡, 四川 643000, 中国  
Institution Address:  
宜宾学院三江人工智能与机器人研究院, 宜宾, 四川 644000, 中国  
Institution Address:  
School of Automation and Information Engineering,Sichuan University of Science & Engineering, Zigong, Sichuan 643000, China  
Institution Address:  
Sanjiang Artificial Intelligence and Robot Research Institute,Yibin University, Yibin, Sichuan 644000, China  
Institution:  
四川轻化工大学自动化与信息工程学院  
Institution:  
宜宾学院三江人工智能与机器人研究院  
Institution:  
School of Automation and Information Engineering,Sichuan University of Science & Engineering  
Institution:  
Sanjiang Artificial Intelligence and Robot Research Institute,Yibin University

### 682. MMFL-net: multi-scale and multi-granularity feature learning for

摘要: Instance-level image retrieval in fashion industry is a challenging issue owing to its increasing importance in real-scenario visual fashion search. Cross-domain fashion retrieval aims to match the unconstrained customer images as queries for photographs provided by retailers; however, it is a difficult task due to a wide range of consumer-to-shop (C2S) domain discrepancies and also considering that clothing image is vulnerable to various non-rigid deformations. To this end, we propose a novel multi-scale and multi-granularity feature learning network (MMFL-net), which can jointly learn global-local aggregation feature representations of clothing images in a unified framework, aiming to train a cross-domain model for C2S fashion visual similarity. First, a new semantic-spatial feature fusion part is designed to bridge the semantic-spatial gap by applying top-down and bottom-up bidirectional multi-scale feature fusion. Next, a multi-branch deep network architecture is introduced to capture global salient, part-informed, and local detailed information, and extracting robust and discrimination feature embedding by integrating the similarity learning of coarse-to-fine embedding with the multiple granularities. Finally, the improved trihard loss, center loss, and multi-task classification loss are adopted for our MMFL-net, which can jointly optimize intra-class and inter-class distance and thus explicitly improve intra-class compactness and inter-class discriminability between its visual representations for feature learning. Furthermore, our proposed model also combines the multi-task attribute recognition and classification module with multi-label semantic attributes and product ID labels. Experimental results demonstrate that our proposed MMFL-net achieves significant improvement over the state-of-the-art methods on the two datasets, DeepFashion-C2S and Street2Shop. Specifically, our approach exceeds the current best method by a large margin of +4.2% and + 11.4% for mAP and Acc@1, respectively, on the most challenging dataset DeepFashion-C2S.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:000842923700001  
Indexed Date:  
2022-08-30  
Institution Address:  
Zhejiang Univ Technol, Coll Comp Sci & Technol, Hangzhou, Peoples R China  
Institution Address:  
Zhejiang Shuren Univ, Sch Informat Sci & Technol, Hangzhou, Peoples R China  
Institution Address:  
Hangzhou Normal Univ, Sch Informat Sci & Technol, Hangzhou, Peoples R China  
Institution:  
Institution:  
Institution:

### 683. AMR-Net: Arbitrary-Oriented Ship Detection Using Attention Module,

摘要: Ship detection is significant and full of challenges in the field of remote sensing. The widely adopted horizontal bounding box representation is not appropriate for ubiquitous oriented ship objects. Complex backgrounds, small objects, and the dilemma of labeling ship datasets are all obstacles that further limit the successful operation of traditional methods in ship detection. In this paper, we propose a multi-task rotation detector using attention module, multi-scale feature fusion and rotation pseudo-label, namely AMR-Net. AMR-Net adds Deformable Convolution Channel Attention Block (DCCAB) to suppress background noise and highlight the foreground. Feature Pyramid Network (FPN) fuses features from different scales, which is beneficial for ship detection. We present Adaptive FPN (AFPN) to determine the optimal number of prediction layers automatically, which can reduce the disturbance of high-level detection of small objects, decrease the size of the model, and adapt to different ship datasets. To enable the detector to achieve excellent performance even when using fewer labeled data, a semi-supervised pseudo-label module is designed, namely Self-Learning Rotation Pseudo-Label (SRP). SRP allows the detector to iteratively self-learn the optimal thresholds, and use the thresholds to refine high-quality rotation pseudo-labels for retraining the model. SRP is optional. If this module is selected, the detector becomes a semi-supervised detector. Extensive both supervised and semi-supervised experiments on remote sensing public datasets HRSC2016 show the state-of-the-art performance of our detector. Experiments on DOTA further illustrate the effectiveness of AMR-Net.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
6  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
被引频次合计:  
9  
ISSN:  
2169-3536  
入藏号:  
WOS:000649566900001  
Indexed Date:  
2021-05-19  
Institution Address:  
Beihang Univ, Sch Elect & Informat Engn, Beijing 100191, Peoples R China  
Institution Address:  
Beijing Inst Control & Elect Technol, Beijing 100038, Peoples R China  
Institution:  
Institution:  
Beijing Inst Control & Elect Technol

### 684. Enhancing medical text detection with vision-language pre-training and

摘要: Detecting text within medical images presents a formidable challenge in the domain of computer vision due to the intricate nature of textual backgrounds, the dense text concentration, and the possible existence of extreme aspect ratios. This paper introduces an effective and precise text detection system tailored to address these challenges. The system incorporates an optimized segmentation module, a trainable post-processing method, and leverages a vision-language pre-training model (oCLIP). Specifically, our segmentation head integrates three essential components: the Feature Pyramid Network (FPN) module, which combines a residual structure and channel attention mechanism; the Efficient Feature Enhancement Module (EFEM); and the Multi-Scale Feature Fusion with RSEConv (MSFM-RSE), designed specifically for multi-scale feature fusion based on RSEConv. By introducing a residual structure and channel attention mechanism into the FPN module, the convolutional layers are replaced with RSEConv layers that employ a channel attention mechanism, further augmenting the representational capacity of the feature maps. The EFEM, designed as a cascaded U-shaped module, incorporates a spatial attention mechanism to introduce multi-level information, thereby enhancing segmentation performance. Subsequently, the MSFM-RSE adeptly amalgamates features from various depths and scales of the EFEM to generate comprehensive final features tailored for segmentation purposes. Additionally, a post-processing module employs a differentiable binarization strategy, allowing the segmentation network to dynamically determine the binarization threshold. Building on the system's improvement, we introduce a vision-language pre-training model that undergoes extensive training on various visual language understanding tasks. This pre-trained model acquires detailed visual and semantic representations, further reinforcing both the accuracy and robustness in text detection when integrated with the segmentation module. The performance of our proposed model was evaluated through experiments on medical text image datasets, demonstrating excellent results. Multiple benchmark experiments validate its superior performance in comparison to existing methods. Code is available at: https://github.com/csworkcode/VLDBNet.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
2199-4536  
eISSN:  
2198-6053  
入藏号:  
WOS:001172827900008  
Indexed Date:  
2024-04-24  
Institution Address:  
Northeast Elect Power Univ, Coll Comp Sci & Technol, Jilin 132012, Peoples R China  
Institution Address:  
Jiangxi New Energy Technol Inst, Xinyu 33800, Jiangxi, Peoples R China  
Institution:  
Institution:  
Jiangxi New Energy Technol Inst

### 685. Two-Layer Attention Feature Pyramid Network for Small Object Detection

摘要: Effective small object detection is crucial in various applications including urban intelligent transportation and pedestrian detection. However, small objects are difficult to detect accurately because they contain less information. Many current methods, particularly those based on Feature Pyramid Network (FPN), address this challenge by leveraging multi-scale feature fusion. However, existing FPN-based methods often suffer from inadequate feature fusion due to varying resolutions across different layers, leading to suboptimal small object detection. To address this problem, we propose the Two-layer Attention Feature Pyramid Network (TA-FPN), featuring two key modules: the Two-layer Attention Module (TAM) and the Small Object Detail Enhancement Module (SODEM). TAM uses the attention module to make the network more focused on the semantic information of the object and fuse it to the lower layer, so that each layer contains similar semantic information, to alleviate the problem of small object information being submerged due to semantic gaps between different layers. At the same time, SODEM is introduced to strengthen the local features of the object, suppress background noise, enhance the information details of the small object, and fuse the enhanced features to other feature layers to ensure that each layer is rich in small object information, to improve small object detection accuracy. Our extensive experiments on challenging datasets such as Microsoft Common Objects in Context (MS COCO) and Pattern Analysis Statistical Modelling and Computational Learning, Visual Object Classes (PASCAL VOC) demonstrate the validity of the proposed method. Experimental results show a significant improvement in small object detection accuracy compared to state-of-theart detectors.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1526-1492  
eISSN:  
1526-1506  
入藏号:  
WOS:001266307500001  
Indexed Date:  
2024-07-16  
Institution Address:  
Zhejiang Univ Technol, Coll Informat Engn, Hangzhou 310023, Peoples R China  
Institution Address:  
ZJUT, Binjiang Cyberspace Secur Inst, Hangzhou 310056, Peoples R China  
Institution:  
Institution:  
ZJUT

### 686. Gesture Recognition of Traffic Police Based on Spatio-Temporal Feature

摘要: In recent years, with the development of human pose estimation technology, gesture recognition technology based on skeleton key points comes into being. This paper proposes a GCPM-AGRU model for gesture recognition of traffic police. In order to locate the key points of human body more accurately, the convolution pose machine(CPM) is improved. Firstly, the idea of residuals, channel split and channel shuffle are added to the feature extraction module, so that the designed feature extraction module can better extract image features. In addition, the parallel multi-branch Inception4d structure is added in the first stage of CPM, which makes the CPM network have the idea of multi-scale feature fusion, and effectively improves the problem of human key point location. Secondly, a GRU based on attention mechanism is proposed, which allocates different weights to each frame to achieve different degrees of attention to each frame, so as to obtain better time information. Finally, it combines the spatio-temporal feature information to carry out traffic police gesture recognition. The accuracy of traffic police gesture recognition reaches 93.7%, which is 2.95 percentage points higher than before the improvement of network.  
摘要:  
近年来,随着人体姿态估计技术的发展,基于骨架关键点的手势识别技术应运而生。提出了一个GCPM-AGRU模型进行交通警察手势识别。为了更准确地定位人体关键点,对卷积姿态机(CPM)进行改进。在特征提取模块中加入残差思想、通道拆分和通道重组,设计后的特征提取模块更好提取图片特征;在CPM第一阶段加入并行多分支Inception4d结构,使CPM网络具有多尺度特征融合思想,有效改进对人体关键点定位的问题;提出基于注意力机制的GRU,通过为每帧分配不同权重来达到对每帧不同程度的关注,从而更好获取时间信息;结合时空特征信息进行交通警察手势识别。交通警察手势识别的准确度达到了93.7%,相比网络改进之前提高了2.95个百分点。  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7708110  
Indexed Date:  
2024-07-12  
Institution Address:  
辽宁科技大学计算机与软件工程学院, 鞍山, 辽宁 114051, 中国  
Institution Address:  
School of Computer and Software Engineering, Liaoning University of Science and Technology, Anshan, Liaoning 114051, China  
Institution:  
辽宁科技大学计算机与软件工程学院  
Institution:  
School of Computer and Software Engineering, Liaoning University of Science and Technology

### 687. Detection of Minor Defects on the Surface of Hydraulic Valve Block Based

摘要: To address issues such as the extremely small size of hydraulic valve block surface defects,low contrast, and significant surrounding interference information,which lead to a high leakage detection rate and low recognition accuracy,a detection algorithm of minor defects on the surface of a hydraulic valve block based on improved YOLOv7 algorithm is proposed. First,a CA attention mechanism is introduced after the multi-scale feature fusion module to improve attention to the feature information of minor defects. Then,the improved UpC multi-branch upsampling structure is used to replace the nearest-neighbor interpolation UpSampling in the multi-scale feature fusion module to enrich the feature information of minor defects. Finally,an improved ELAN-RepConv structure is used to replace the ELAN\_2 structure in the multi-scale feature fusion module,so that the model can learn more feature information during the training process. To improve the robustness and convergence speed of the algorithm further,offline data augmentation,fusing Mosaic data augmentation,and the K-means++ clustering anchor box algorithm are used to enhance the performance of the algorithm. The experimental results indicate that the Average Precision(AP) value of this algorithm on the dataset of minor defects on the surface of the hydraulic valve block is 97.6%,8.4 percentage points higher than the original YOLOv7 algorithm,and the detection speed reaches 55.2 frame/s. Compared with the YOLOv7- E6E algorithm,which has the highest detection accuracy in the YOLOv7 series,the AP value is improved by 1.8 percentage points when the number of parameters is reduced by 75.4%. The experimental results show that the improved algorithm can improve detection precision on the premise of ensuring real-time.  
摘要:  
针对液压阀块表面缺陷尺寸微小、对比度低、周围干扰信息多导致的漏检率高、识别准确率低等问题,提出一种基于改进YOLOv7的液压阀块表面微小缺陷检测算法。在多尺度特征融合模块后引入CA注意力机制来提高对微小缺陷特征信息的关注度。使用改进的UpC多支路上采样结构代替多尺度特征融合模块中的最近邻插值上采样UpSampling模块,以丰富微小缺陷的特征信息。利用改进的ELAN-RepConv结构代替多尺度特征融合模块中的ELAN\_2结构,使模型在训练过程中可以学习到更多的特征信息。为了进一步提高算法的鲁棒性与收敛速度,使用离线数据增强融合Mosaic数据增强的数据增广技术与K-means++锚框聚类算法来提高算法性能。实验结果表明:该算法在液压阀块表面微小缺陷数据集中平均精度达到97.6%,较原YOLOv7算法提高8.4个百分点,检测速度达到55.2 frame/s;相较于YOLOv7系列中检测精度最高的YOLOv7-E6E算法,该算法在参数量减少75.4%的情况下,平均精度值提高1.8个百分点。所提算法在保证实时性的前提下能够有效提高检测精度。  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1000-3428  
入藏号:  
CSCD:7593415  
Indexed Date:  
2024-01-19  
Institution Address:  
江苏科技大学机械工程学院, 镇江, 江苏 212100, 中国  
Institution Address:  
中国船舶重工集团公司上海船舶设备研究所, 上海 200031, 中国  
Institution Address:  
江苏科技大学机械工程学院;;江苏科技大学海洋装备研究院, ;;, 镇江;;镇江, 江苏;;江苏 212100;;212003, 中国  
Institution Address:  
School of Mechanical Engineering,Jiangsu University of Science and Technology, Zhenjiang, Jiangsu 212100, China  
Institution Address:  
Shanghai Marine Equipment Research Institute,China Shipbuilding Industry Group Co.,Ltd., Shanghai 200031, China  
Institution Address:  
School of Mechanical Engineering,Jiangsu University of Science and Technology;;Marine Equipment and Technology Institute,Jiangsu University of Science and Technology, ;;, Zhenjiang;;Zhenjiang, Jiangsu;;Jiangsu 212100;;212003  
Institution:  
江苏科技大学机械工程学院  
Institution:  
中国船舶重工集团公司上海船舶设备研究所  
Institution:  
江苏科技大学机械工程学院;;江苏科技大学海洋装备研究院  
Institution:  
School of Mechanical Engineering,Jiangsu University of Science and Technology  
Institution:  
Shanghai Marine Equipment Research Institute,China Shipbuilding Industry Group Co.,Ltd.  
Institution:  
School of Mechanical Engineering,Jiangsu University of Science and Technology;;Marine Equipment and Technology Institute,Jiangsu University of Science and Technology

### 688. Research on Railway Track Foreign Object Intrusion Detection Based on

摘要: To address the issues of low detection accuracy and slow detection speed caused by the influence of complex environments on detecting foreign objects of different scales on railroad tracks, this paper proposes a multi-scale feature fusion RMF-YOLO(railway multi-scale fusion YOLO)method for railroad track foreign object detection. Firstly, ICBAM, an improved convolutional attention module, is designed and introduced in combination with the YOLOv7 feature extraction network to enhance feature extraction capability in complex scenarios. Subsequently, GhostConv is adopted in all efficient layer aggregation network modules instead of regular convolutional layers to reduce computational complexity and enhance feature output efficiency.An improved weighted bi-directional feature pyramid network(N-BiFPN)structure is introduced to enhance multi-scale feature fusion, balance feature information across different levels, and improve multi-scale detection capabilities. Lastly, to further enhance detection accuracy, the WIoU loss function combined with the dynamic non-monotonic aggregation mechanism is used to effectively deal with the gradient generated by the low-quality anchor frames and to improve the overall performance of the model for the detection of foreign objects at different scales. The experimental results indicate that on the self-made railway track foreign object dataset, the RMF-YOLO algorithm reduces the parameter count of the original network model, effectively enhancing the model's detection accuracy and detection speed of the model and addressing issues related to missed and false detections. On average, the precision increased by 5.5%, the frames per second(FPS)improved by 5.88%, and the computational load decreased by 12.25%. These enhancements meet the requirements for both detection accuracy and real-time performance in railway track foreign object detection.  
摘要:  
针对铁路轨道异物检测中不同尺度异物目标的检测易受复杂环境的影响,导致出现检测精度低及检测速度慢等问题,提出一种多尺度特征融合的铁轨异物检测(RMF-YOLO)算法。首先,设计并引入改进的卷积注意力模块(ICBAM),结合YOLOv7特征提取网络,以增强复杂场景下的特征提取能力。其次,在所有高效层聚合网络模块中采用GhostConv替代普通卷积层,以降低计算复杂度,提高特征输出效率;设计一种改进的加权双向特征金字塔网络N-BiFPN结构,加强多尺度特征融合能力,平衡不同层级特征信息,提高多尺度检测能力。最后,为进一步提升检测精度,采用WIoU损失函数结合动态非单调聚集机制,有效应对低质量锚框产生的梯度,提高模型对不同尺度异物检测的整体性能。实验结果表明:在自制的铁轨异物数据集上,RMF-YOLO算法减少了原网络模型的参数量,有效提升了模型的检测精度与检测速度,改善了漏检与误检问题,平均精度提升了5.5%,检测速度提升了5.88%,计算量减少了12.25%,能满足铁轨入侵异物检测中对检测精度和实时性的需求。  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0253-987X  
入藏号:  
CSCD:7818195  
Indexed Date:  
2024-12-21  
Institution Address:  
兰州交通大学自动化与电气工程学院, 兰州, 甘肃 730070, 中国  
Institution Address:  
School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China  
Institution:  
兰州交通大学自动化与电气工程学院  
Institution:  
School of Automation and Electrical Engineering, Lanzhou Jiaotong University

### 689. SPA: Leveraging the SAM with Spatial Priors Adapter for Enhanced Medical

摘要: The Segment Anything Model (SAM) has gained renown for its success in image segmentation, benefiting significantly from its pretraining on extensive datasets and its interactive prompt-based segmentation approach. Although highly effective in natural (real-world) image segmentation tasks, the SAM model encounters significant challenges in medical imaging due to the inherent differences between these two domains. To address these challenges, we propose the Spatial Prior Adapter (SPA) scheme, a parameter-efficient fine-tuning strategy that enhances SAM's adaptability to medical imaging tasks. SPA introduces two novel modules: the Spatial Prior Module (SPM), which captures localized spatial features through convolutional layers, and the Feature Communication Module (FCM), which integrates these features into SAM's image encoder via cross-attention mechanisms. Furthermore, we develop a Multiscale Feature Fusion Module (MSFFM) to enhance SAM's end-to-end segmentation capabilities by effectively aggregating multiscale contextual information. These lightweight modules require minimal computational resources while significantly boosting segmentation performance. Our approach demonstrates superior performance in both prompt-based and end-to-end segmentation scenarios through extensive experiments on publicly available medical imaging datasets. Performance highlights the potential of the proposed method to bridge the gap between foundation models and domain-specific medical imaging tasks. This advancement paves the way for more effective AI-assisted medical diagnostic systems.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2168-2208  
入藏号:  
MEDLINE:40030852  
PubMed ID:  
40030852  
Indexed Date:  
2025-03-06

### 690. ICA-Net: Industrial defect detection network based on convolutional

摘要: Detecting surface defects in the industry is essential for improving the quality of industrial products and maintaining product safety. However, problems such as the similarity of defects, significant variation in the scale of the target object, and the balance between detection speed and accuracy in industrial inspection scenarios have been considerable research topics in this field. This paper proposes an industrial defect detection network based on convolutional attention-guided and aggregated multiscale features to address these issues (ICA-Net). Firstly, for similarity defects in complex backgrounds, this paper proposes a backbone network with a combination of lightweight convolutional blocks and self-attentive modules to fully extract images' local and global information and enhance the network's expressiveness. Secondly, to make full use of the shallow fine-grained features and deep semantic features of the backbone network to improve the detection capability of defects with significant scale changes, this paper designs a cross-layer multiscale feature fusion network (CEF-Net), which fully fuses the features of adjacent layers and cross-layers through a reweighting feature strategy to enrich the network feature transfer path and ensure the efficient fusion of different scale features in the network. At the same time, the fine-grained feature fusion module (FFM) is used to fuse elements from multiple layers to extract more contextual information, enhance the extraction of fine-grained features and improve the detection capability of complex small targets. Finally, to address the problems of inaccurate regression localization and low detection accuracy of defects in existing industrial algorithms, a new IoU loss function (G-IOU) is proposed for regressing the intersection part of the predicted frame and the actual structure according to the aspect ratio of the real frame during the model regression to improve the accuracy and stability of detection. The experimental results show that 94.1%, 98.6%, 99.4%, 98.8% and 96.5% of mAP@.5 are obtained on steel, PCB, aluminium, automobile and Xsteel steel metal surface defect datasets, respectively, and 48 FPS is achieved, which is superior to the current mainstream detectors and meets the needs of practical industrial production.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001106151900001  
Indexed Date:  
2023-12-04  
Institution Address:  
Qilu Univ Technol, Shandong Comp Sci Ctr, Natl Supercomp Ctr Jinan, Shandong Acad Sci,Key Lab Comp Power Network & Inf, Jinan 250353, Peoples R China  
Institution:

### 691. A Multiscale Recursive Attention Gate Federation Method for Multiple

摘要: Federated learning (FL) is an effective method when a single client cannot provide enough samples for multiple condition fault diagnosis of bearings since it can combine the information provided by multiple clients. However, some of the client's working conditions are different; for example, different clients are in different stages of the whole life cycle, and different clients have different loads. At this point, the status of each client is not equal, and the traditional FL approach will lead to some clients' useful information being ignored. The purpose of this paper is to investigate a multiscale recursive FL framework that makes the server more focused on the useful information provided by the clients to ensure the effectiveness of FL. The proposed FL method can build reliable multiple working condition fault diagnosis models due to the increased focus on useful information in the FL process and the full utilization of server information through local multiscale feature fusion. The validity of the proposed method was verified with the Case Western Reserve University benchmark dataset. With less local client training data and complex fault types, the proposed method improves the accuracy of fault diagnosis by 23.21% over the existing FL fault diagnosis.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
1099-4300  
入藏号:  
WOS:001057029200001  
PubMed ID:  
37628195  
Indexed Date:  
2023-09-12  
Institution Address:  
Shanghai Maritime Univ, Sch Logist Engn, Shanghai 201306, Peoples R China  
Institution Address:  
Chinese Acad Sci, Shenzhen Inst Adv Technol, Guangdong Prov Key Lab Robot & Intelligent Syst, Shenzhen 518055, Peoples R China  
Institution Address:  
Guangdong Univ Petrochem Technol, Sch Automat, Maoming 525000, Peoples R China  
Institution:  
Institution:  
Institution:

### 692. An efficient multiscale enhancement network with attention mechanism for

摘要: Real-time defect detection is required to efficiently control the quality of aluminium. However, aluminium defects have the characteristics of small-size, low contrast, and multiscale variations, which pose great challenges to defect detection. This article aims to improve the defect detection accuracy and propose an effective multiscale enhancement network with attention mechanism for aluminium defect detection (AMMENet). First, to capture key features and mitigate interference from the background, a pluggable parallel residual attention module (PRAM) is proposed for the feature extraction network. To compensate for the loss of deep features, a multilevel semantic enhancement module (C2f-MFF) is proposed to fuse multiscale feature maps. Finally, the model was applied to the Tianchi Aluminium Surface Defect Dataset (TC-ASDD) for ablation experiments and comparisons. The experimental results show that the mean average precision (mAP@0.5) of the proposed AMMENet is 73.6% with a real-time detection speed of 66.2 frames per second (FPS). Compared with YOLOv8 baseline network, AMMENet improves mAP@0.5 by 2.8% with only a slight loss in speed. Moreover, AMMENet is superior to the state-of-the-art detection methods in terms of detection accuracy.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1058-9759  
eISSN:  
1477-2671  
入藏号:  
WOS:001336733800001  
Indexed Date:  
2024-10-27  
Institution Address:  
Shanghai DianJi Univ, Sch Elect Informat Engn, Shanghai, Peoples R China  
Institution:

### 693. An Efficient Bidirectional Point Pyramid Attention Network for 3D Point

摘要: Point cloud completion is a necessary task in real-world applications of recovering a complete geometry from missing regions of 3D objects. Furthermore, model efficiency is of vital importance in computer vision. In this paper, we present an efficient encoder-decoder network that predicts missing point clouds on the basis of incomplete point clouds. There are several advantages to this approach. First, a Mixed Attention Module (MAM) was implemented to obtain the correlational information of points. Second, the proposed Bidirectional Point Pyramid Attention Network (BiPPAN) can achieve simple and fast multiscale feature fusion to capture important features. Lastly, the designed encoder-decoder framework comprises skip connections to capture long-distance dependencies and structural information. We can conclude from the results of the experiments that the proposed network is an efficient and effective method to accomplish point cloud completion tasks.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:000979104500001  
Indexed Date:  
2023-05-15  
Institution Address:  
Dalian Neusoft Univ Informat, Sch Digital Arts & Design, Dalian 116023, Peoples R China  
Institution:  
Dalian Neusoft Univ Informat

### 694. Enhancing instance segmentation: Leveraging multiscale feature fusion

摘要: The accurate estimation of fish weight relies on the crucial parameter of individual fish contour features. While instance segmentation proves effective in extracting fish contours, challenges arise from diverse fish postures and reduced image sharpness underwater. Current instance segmentation methods often struggle to effectively balance global and local detailed features, which can result in inaccurate positioning of contour keypoints and consequently limit the accuracy of fish weight estimation. To overcome this, our study introduces a novel instance segmentation network tailored for precise fish contour extraction. The proposed approach incorporates multi-scale feature fusion and an attention mechanism based on the Segmenting Objects by Locations (SOLO) network, referred to as SOLO-MFFA. This paper designs a multi-scale context aggregation module to integrate features with a wider range of receptive fields, augmenting the model's capability to comprehend both local features and global information. At the same time, the introduction of a mixed-domain attention mechanism emphasizes more critical channel features and simultaneously improves the localization accuracy of contour points. Compared with SOLO and its improved model CAM-SOLO on the fish instance segmentation dataset, SOLO-MFFA demonstrated an effective improvement, with a 4.3% and 1.6% increase in mAP (mean Average Precision), respectively. The Decoupled-SOLO-MFFA achieved higher mAP. The visualization results also demonstrate that the contour features extracted in this paper are smoother and more accurately positioned. Additionally, in comparison to other well-known instance segmentation networks, our method has demonstrated significant improvements in both qualitative and quantitative evaluations. Furthermore, the integration of contour features derived from Decoupled-SOLO-MFFA, along with binocular vision, was utilized for the precise estimation of fish perimeter and weight. The findings reveal a strong correlation between the perimeter calculated by Decoupled-SOLO-MFFA and the actual weight, with a notably reduced error in weight estimation. Compared to previous methods, RMSE, MAE, and MAPE of the linear model constructed in this paper decreased by 3.92, 3.19, and 1.4%.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0144-8609  
eISSN:  
1873-5614  
入藏号:  
WOS:001244321000001  
Indexed Date:  
2024-06-17  
Institution Address:  
China Agr Univ, Coll Informat & Elect Engn, Beijing 100083, Peoples R China  
Institution Address:  
Yantai Univ, Sch Comp & Control Engn, Yantai 264005, Peoples R China  
Institution Address:  
Shandong Technol & Business Univ, Sch Comp Sci & Technol, Yantai 264005, Peoples R China  
Institution Address:  
Qingdao Agr Univ, Coll Sci & Informat, Qingdao 266109, Peoples R China  
Institution Address:  
China Agr Univ, Natl Innovat Ctr Digital Fishery, Beijing 100083, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 695. A Ship Detection Method from Lightweight SAR Images under Complex

摘要: The ship target detection methods using Synthetic Aperture Radar (SAR) images have a wide range of practical applications in many fields such as surveillance on the sea surface, trade to and from the sea surface, and emergency rescue on the sea surface, etc. With the demand for the development of autonomous processing in satellite orbits, the real-time in-orbit detection and localization of ships from SAR images have put forward higher requirements. Therefore, this paper proposes a lightweight SAR image ship detection algorithm in a complex background for the current problems of limited satellite hardware resources, diverse and differentiated feature scales of different ship targets in Synthetic Aperture Radar (SAR) images, and easy to be interfered by noise. First of all, the FasterNet network model combined with the attention mechanism is used to extract different high and low level features of the target. Second, in order to solve the problem of scale inconsistency between different targets, this paper constructs a Feature Enhancement Module (FEM) that can not only increase the network sensory field at the same time but also improve the ability of network target detection. Then, a multiscale feature fusion structure combined with feature enhancement is constructed in this paper, which can enhance and fuse the multi-scale features extracted by the backbone feature extraction network, and can also strengthen the connection between the features of different layers of the network while obtaining the multi-scale contextual information of the target, and carry out the detection of the SAR image ship in the three feature maps output from the multi-scale feature fusion structure combined with feature enhancement. Experiments are conducted to compare the proposed method with some other mainstream target detection algorithms on SSDD, HRSID, and merged SSDD and HRSID datasets. The results show that the average accuracy of the proposed methods on three datasets in this paper is 98.6%, 92.3% and 93.0%, respectively. The recall of the method in this paper is 95.10%, 85.10% and 86.8%, respectively, for three datasets. The model size and parameter number of the proposed method in this paper are only 8.8 MB and 4.2 M, respectively. The proposed method significantly outperforms other algorithms in terms of recall and average accuracy ratio. Moreover, the method in this paper also has great advantages in terms of checking accuracy and detection rate, which is favorable to be migrated to other practical applications.  
摘要:  
合成孔径雷达(SAR)影像船舶检测在海上监视、海上贸易和海上救援等领域有着广泛的实际应用价值,随着星上在轨自主处理发展的需求,对星上SAR影像船舶在轨实时检测提出了更高的要求。因此针对当前卫星硬件资源有限、SAR影像船舶目标特征尺度多样化、易受噪声干扰等问题,本文提出了一种复杂背景下轻量级SAR影像船舶检测算法。首先,采用FasterNet结合注意力机制提取目标的高低级特征;其次,为解决目标多尺度问题,构建了一种特征增强模块(FEM),在增加感受野的同时,提高了目标检测能力;然后,构造结合特征增强的多尺度特征融合结构,将主干网络提取的多尺度特征进行增强和融合,加强不同层级特征联系的同时获取目标多尺度的上下文信息,并在输出特征图中进行SAR影像船舶检测;最后,在SSDD、HRSID以及二者合并的数据集上,将本文方法与主流算法进行对比实验。实验结果表明,本文方法平均准确率分别为98.6%、92.3%和93.0%,召回率分别为95.10%、85.10%和86.8%,模型尺寸和参数量仅为8.8 MB和4.2 M,均明显优于其他方法,并且查准率和检测速率也具有很大的优势,有利于迁移到实际的应用中。  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1560-8999  
入藏号:  
CSCD:7847046  
Indexed Date:  
2025-01-25  
Institution Address:  
中国人民解放军战略支援部队信息工程大学, 郑州, 河南 450001, 中国  
Institution Address:  
山东五征集团, 日照, 山东 262300, 中国  
Institution Address:  
PLA Strategic Support Force Information Engineering University, Zhengzhou, Henan 450001, China  
Institution Address:  
Shangdong Wuzheng Group Co.,Ltd., Rizhao, Shandong 262300, China  
Institution:  
中国人民解放军战略支援部队信息工程大学  
Institution:  
山东五征集团  
Institution:  
PLA Strategic Support Force Information Engineering University  
Institution:  
Shangdong Wuzheng Group Co.,Ltd.

### 696. Accurate and real-time visual detection algorithm for environmental

摘要: Owing to the intricate and ever-changing nature of the marine environment, traditional marine survey methods are subject to numerous limitations. Unmanned surface vehicles (USVs) have gained significant popularity for their role in automatically identifying and positioning targets in the ocean. To enhance the environmental perception capabilities of USVs in complex marine environments, vision-based sea surface object detection algorithms have emerged as a crucial technological approach. In response to the unique challenges of sea surface object detection tasks and the various extreme situations encountered by USVs in real sea environments, YOLOv7 was chosen as the baseline model due to its excellent trade-off between speed and accuracy. We propose Efficient Multi-Scale Pyramid Attention Networks, which enable easy and rapid multi-scale feature fusion. Additionally, we improve the boundary box loss function. Building upon these optimizations, we develop a new detector called Marit-YOLO, which consistently achieves better accuracy and efficiency than the prior art across a wide spectrum of resource constraints. Marit-YOLO achieved a 5.0% increase in the average precision (AP) value on the independently collected Ocean Buoys Dataset. During generalization validation, the AP value also saw a 7.7% increase on the open-source Singapore Maritime Dataset. The Marit-YOLO algorithm achieves a real-time inference speed of 69 frames per second on a single RTX2080, enabling accurate and real-time target detection in complex sea environments.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001169638800001  
Indexed Date:  
2024-03-09  
Institution Address:  
Harbin Engn Univ, Coll Intelligent Syst Sci & Engn, 145 Nantong St, Harbin 150000, Heilongjiang, Peoples R China  
Institution:

### 697. Feature Fusion Target Detection Algorithm Using Dynamic Sample

摘要: A multi-scale feature fusion target detection algorithm with dynamic sample allocation strategy is proposed to address the problems of low detection accuracy and poor prediction ability of small targets in the lightweight target detection algorithm SSD-Lite. Firstly, the feature pyramid network (FPN) is introduced in the neck network of the lightweight target detection algorithm SSD-Lite and designed to be lightweight, while the residual feature augmentation (RFA) module is introduced, which uses residual branches to inject different. Then, this paper inserts a lightweight attention mechanism ECA module into the feature pyramid structure to improve the ability of network to focus on important features. Finally, to address the problems of poor adaptability of positive and negative sample assignment and difficulty in selecting high-quality positive samples caused by the fixed Intersection-over-Union (IOU) threshold sample assignment strategy used in the network training process, this paper designs a dynamic sample assignment strategy, which eliminates the pre-setting of anchor frames and adopts the centroid sampling method, while combining the sample mean and standard deviation as screening thresholds to reduce the influence of artificial a priori and improve the algorithm performance without changing the network structure. The algorithm is tested on Pascal VOC dataset, and the experimental results show that the overall prediction accuracy of the algorithm is improved by 1.9 percentage points compared with the benchmark algorithm, the detection ability of small targets is improved by 3.3 percentage points, and the inference delay of the algorithm is increased by only 2.32%. The experiments demonstrate that the algorithm can significantly improve the prediction accuracy of the algorithm with a small performance cost.  
摘要:  
针对轻量级目标检测算法SSD-Lite检测精度低、对小目标预测能力差等问题,提出了一种采用动态样本分配策略的多尺度特征融合目标检测算法。在轻量级目标检测算法SSD-Lite的颈部网络引入特征金字塔结构(feature pyramid network,FPN),并对其进行轻量化设计,同时引入残差特征增强模块(residual feature augmentation,RFA),采用残差分支注入不同空间的上下文信息来改善高层特征的特征表达,以提升网络对小目标的检测能力;在特征金字塔结构中插入轻量级注意力机制ECA模块,提升网络对重要特征的关注能力;针对网络训练过程中采用的固定交并比(intersection-over-union,IOU)阈值的样本分配策略导致的正负样本分配适应性差、难以选出高质量正样本等问题,设计了一种动态样本分配策略,取消锚框的预设置,采用中心点采样的方式,同时结合样本均值、标准差作为筛选阈值,减少人工先验的影响,在不改变网络结构的情况下提升算法性能。算法在Pascal VOC数据集上测试,实验结果表明:该算法整体预测精度相较于基准算法提升1.9个百分点,对小目标检测能力提升3.3个百分点,算法推理时延仅增加2.32%;实验证明了该算法可以以较小的性能代价,显著提升算法的预测精度。  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7804296  
Indexed Date:  
2024-11-23  
Institution Address:  
西安工业大学电子信息工程学院, 西安, 陕西 710021, 中国  
Institution Address:  
西安工业大学发展规划处, 西安, 陕西 710021, 中国  
Institution Address:  
西安机电信息技术研究所, 西安, 陕西 710065, 中国  
Institution Address:  
西安工业大学光电工程学院, 西安, 陕西 710021, 中国  
Institution Address:  
School of Electronic Information Engineering,Xi'an Technological University, Xi'an, Shaanxi 710021, China  
Institution Address:  
Development Planning Service,Xi'an Technological University, Xi'an, Shaanxi 710021, China  
Institution Address:  
Xi'an Institute of Mechanical and Electrical Information Technology, Xi'an, Shaanxi 710065, China  
Institution Address:  
School of Optoelectronic Engineering,Xi'an Technological University, Xi'an, Shaanxi 710021, China  
Institution:  
西安工业大学电子信息工程学院  
Institution:  
西安工业大学发展规划处  
Institution:  
西安机电信息技术研究所  
Institution:  
西安工业大学光电工程学院  
Institution:  
School of Electronic Information Engineering,Xi'an Technological University  
Institution:  
Development Planning Service,Xi'an Technological University  
Institution:  
Xi'an Institute of Mechanical and Electrical Information Technology  
Institution:  
School of Optoelectronic Engineering,Xi'an Technological University

### 698. Fine-Grained Semantic Image Synthesis with Object-Attention Generative

摘要: Semantic image synthesis is a new rising and challenging vision problem accompanied by the recent promising advances in generative adversarial networks. The existing semantic image synthesis methods only consider the global information provided by the semantic segmentation mask, such as class label, global layout, and location, so the generative models cannot capture the rich local fine-grained information of the images (e.g., object structure, contour, and texture). To address this issue, we adopt a multi-scale feature fusion algorithm to refine the generated images by learning the fine-grained information of the local objects. We propose OA-GAN, a novel object-attention generative adversarial network that allows attention-driven, multi-fusion refinement for fine-grained semantic image synthesis. Specifically, the proposed model first generates multi-scale global image features and local object features, respectively, then the local object features are fused into the global image features to improve the correlation between the local and the global. In the process of feature fusion, the global image features and the local object features are fused through the channel-spatial-wise fusion block to learn 'what' and 'where' to attend in the channel and spatial axes, respectively. The fused features are used to construct correlation filters to obtain feature response maps to determine the locations, contours, and textures of the objects. Extensive quantitative and qualitative experiments on COCO-Stuff, ADE20K and Cityscapes datasets demonstrate that our OA-GAN significantly outperforms the state-of-the-art methods.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2157-6904  
eISSN:  
2157-6912  
入藏号:  
WOS:000732997200009  
Indexed Date:  
2021-12-31  
Institution Address:  
Beijing Jiaotong Univ, Beijing, Peoples R China  
Institution:

### 699. Pointer generation and main scale detection for occluded meter reading

摘要: The meter reading with machine vision greatly improves the efficiency of industrial monitoring. However, the pointer and scales of the meter can be occluded by rain or dirt, which greatly reduces the accuracy of the meter reading recognition. To solve this problem, we propose a generative adversarial network (PMS-GAN) with pointer generation and main scale detection for occluded meter reading. Specifically, dilated convolution block is designed to correlate separated pointer features. Then multi-scale feature fusion mechanism is proposed to guarantee the precision of pointer generation and main scale detection with guidance of semantic information. Moreover, feature enhancement mechanism is proposed to construct the long -range relationship for generating pointer under high occlusion. Finally, the reading is accomplished by calculating local angle with generated pointer and detected main scales. Experiments show that PMS-GAN can generate more intact pointer and detect main scales to guarantee the success and accuracy of occluded meter reading.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
4  
ISSN:  
0263-2241  
eISSN:  
1873-412X  
入藏号:  
WOS:001241823800001  
Indexed Date:  
2024-06-14  
Institution Address:  
Fuzhou Univ, Coll Elect Engn & Automat, Fuzhou 350108, Peoples R China  
Institution Address:  
Yanshan Univ, Sch Elect Engn, Qinhuangdao 066004, Peoples R China  
Institution:  
Institution:

### 700. Multi-scale High-Resolution Preserving and Perspective-Invariant Hand

摘要: At present, most of the networks used for 2D keypoint heatmaps estimation of hand pose use the convolutional pose machines or Hourglass network, but these two networks cannot simultaneously satisfy the requirements of high-resolution representation preserving learning and multi-scale feature fusion.In response to this problem, a multi-scale high-resolution preserving network is used, which adopts the structure of high-resolution and low-resolution representation in parallel design, and enhances the features of each resolution through the fusion of all resolution representations, and has multiple stages to extract high quality features for 2D heatmaps estimation.In order to obtain the 3D hand pose, a global rotation perspective-invariant method is also used to map the 2D heatmaps to the 3D pose.Experiments on 2D hand pose estimation and 3D hand pose estimation are conducted on three public datasets(RHD, STB, Dexter+Object), and the results verify the effectiveness of the method in hand pose estimation.  
摘要:  
目前基于彩色图像的手姿态2D关键点热图估计大多数采用卷积姿势机或沙漏网络进行,但这两种网络不能同时满足高分辨率表示保持学习和多尺度特征融合。针对该问题引用了一种多尺度高分辨率保持的网络,该网络采用高低分辨率表示并行设计的结构,并通过融合所有分辨率表示增强各分辨率表示的特征,而且拥有多个阶段提取高质量特征用于2D热图估计。为得到3D手姿态,还使用了全局旋转视角不变的方法将2D热图映射到3D姿态。在三个公开数据集(RHD、STB、Dexter+Object)上分别对2D手姿态估计和3D手姿态估计进行了实验,结果验证了该方法在手姿态估计中的有效性。  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1002-8331  
入藏号:  
CSCD:7029659  
Indexed Date:  
2021-10-30  
Institution Address:  
江西农业大学计算机与信息工程学院, 南昌, 江西 330045, 中国  
Institution Address:  
江西农业大学软件学院;;浙江大学, ;;CAD&CG国家重点实验室, 南昌;;杭州, ;; 330045;;310058  
Institution Address:  
江铃控股有限公司, 南昌, 江西 330052, 中国  
Institution Address:  
School of Computer and Information Engineering, Jiangxi Agricultural University, Nanchang, Jiangxi 330045, China  
Institution Address:  
School of Software, Jiangxi Agricultural University;;Zhejiang University, ;;State Key Laboratory of CAD&CG, Nanchang;;Hangzhou, ;; 330045;;310058  
Institution Address:  
Jiangling Holdings Limited, Nanchang, Jiangxi 330052, China  
Institution:  
江西农业大学计算机与信息工程学院  
Institution:  
江西农业大学软件学院;;浙江大学  
Institution:  
江铃控股有限公司  
Institution:  
School of Computer and Information Engineering, Jiangxi Agricultural University  
Institution:  
School of Software, Jiangxi Agricultural University;;Zhejiang University  
Institution:  
Jiangling Holdings Limited

### 701. Redefining Contextual and Boundary Synergy: A Boundary-Guided Fusion

摘要: Medical image segmentation plays a crucial role in medical image processing, focusing on the automated extraction of regions of interest (such as organs, lesions, etc.) from medical images. This process supports various clinical applications, including diagnosis, surgical planning, and treatment. In this paper, we introduce a Boundary-guided Context Fusion U-Net (BCF-UNet), a novel approach designed to tackle a critical shortcoming in current methods: the inability to effectively integrate boundary information with semantic context. The BCF-UNet introduces a Adaptive Multi-Frequency Encoder (AMFE), which uses multi-frequency analysis inspired by the Wavelet Transform (WT) to capture both local and global features efficiently. The Adaptive Multi-Frequency Encoder (AMFE) decomposes images into different frequency components and adapts more effectively to boundary texture information through a learnable activation function. Additionally, we introduce a new multi-scale feature fusion module, the Atten-kernel Adaptive Fusion Module (AKAFM), designed to integrate deep semantic information with shallow texture details, significantly bridging the gap between features at different scales. Furthermore, each layer of the encoder sub-network integrates a Boundary-aware Pyramid Module (BAPM), which utilizes a simple and effective method and combines it with a priori knowledge to extract multi-scale edge features to improve the accuracy of boundary segmentation. In BCF-UNet, semantic context is used to guide edge information extraction, enabling the model to more effectively comprehend and identify relationships among various organizational structures. Comprehensive experimental evaluations on two datasets demonstrate that the proposed BCF-UNet achieves superior performance compared to existing state-of-the-art methods.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001387094300001  
Indexed Date:  
2025-01-05  
Institution Address:  
Xiamen Univ Technol, Sch Comp & Informat Engn, Xiamen 361024, Peoples R China  
Institution Address:  
Fujian Key Lab Pattern Recognit & Image Understand, Xiamen 361024, Peoples R China  
Institution:  
Institution:  
Fujian Key Lab Pattern Recognit & Image Understand

### 702. YOLO-LF: a lightweight multi-scale feature fusion algorithm for wheat

摘要: Wheat is one of the most significant crops in China, as its yield directly affects the country's food security. Due to its dense, overlapping, and relatively fuzzy distribution, wheat spikes are prone to being missed in practical detection. Existing object detection models suffer from large model size, high computational complexity, and long computation times. Consequently, this study proposes a lightweight real-time wheat spike detection model called YOLO-LF. Initially, a lightweight backbone network is improved to reduce the model size and lower the number of parameters, thereby improving the runtime speed. Second, the structure of the neck is redesigned in the context of the wheat spike dataset to enhance the feature extraction capability of the network for wheat spikes and to achieve lightweightness. Finally, a lightweight detection head was designed to significantly reduce the FLOPs of the model and achieve further lightweighting. Experimental results on the test set indicate that the size of our model is 1.7 MB, the number of parameters is 0.76 M, and the FLOPs are 2.9, which represent reductions of 73, 74, and 64% compared to YOLOv8n, respectively. Our model demonstrates a latency of 8.6 ms and an FPS of 115 on Titan X, whereas YOLOv8n has a latency of 10.2 ms and an FPS of 97 on the same hardware. In contrast, our model is more lightweight and faster to detect, while the mAP@0.5 only decreases by 0.9%, outperforming YOLOv8 and other mainstream detection networks in overall performance. Consequently, our model can be deployed on mobile devices to provide effective assistance in the real-time detection of wheat spikes.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001286509500001  
Indexed Date:  
2024-08-13  
Institution Address:  
Changsha Univ Sci & Technol, Changsha, Peoples R China  
Institution:

### 703. Multi-Stream Dense View Reconstruction Network for Light Field Image

摘要: Recently, many view synthesis-based methods are proposed for high-efficiency light field (LF) image compression. However, most existing methods fail to recover more texture details on occlusion regions, which reduces the compression efficiency. In this paper, we propose a multi-stream dense view reconstruction network to further improve LF image compression performance. In our method, only sparsely-sampled LF views are transmitted and the rest of the views are reconstructed at the decoder side. During the reconstruction process, we firstly constitute a multi-disparity geometry (MDG) structure based on the decoded sparse LF views, which can reflect abundant disparity characteristics. Subsequently, a multi-stream view reconstruction network (MSVRNet) is put forward to reconstruct a high-quality dense LF image, which consists of a multi-scale feature fusion sub-network, a fusion reconstruction sub-network, and a detail refinement sub-network. The multi-scale feature fusion sub-network can implicitly lean abundant multiscale geometric structure features from the constituted MDG structure. The fusion reconstruction sub-network and the detail refinement sub-network are respectively utilized to fuse the learned multiscale geometric features and restore more texture details, especially for occlusion regions. Moreover, 3D convolutional operations are adopted in the whole reconstruction process, which allow information propagation among the learned multiscale geometric features. Comprehensive experimental results demonstrate the effectiveness of the proposed method. The perceptual quality of reconstructed views and application on depth estimation also demonstrate that the proposed method can keep structural consistency of the reconstructed LF image and recover more texture details.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
15  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
15  
ISSN:  
1520-9210  
eISSN:  
1941-0077  
入藏号:  
WOS:001089390200023  
Indexed Date:  
2023-11-15  
Institution Address:  
Anqing Normal Univ, Anqing 246000, Peoples R China  
Institution Address:  
Jiangxi Univ Finance & Econ, Anqing 246000, Peoples R China  
Institution Address:  
Jiangxi Univ Finance & Econ, Sch Informat Management, Nanchang 330000, Peoples R China  
Institution Address:  
Chinese Acad Sci CASIA, Inst Automat, Ctr Res Intelligent Percept & Comp CRIPAC, Natl Lab Pattern Recognit NLPR, Beijing 100045, Peoples R China  
Institution Address:  
Shanghai Univ, Shanghai Inst Adv Commun & Data Sci, Sch Commun & Informat Engn, Shanghai 200444, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 704. SFINet: A semantic feature interactive learning network for full-time

摘要: Infrared and visible image fusion aims to combine data from various source images to generate a high-quality image. Nevertheless, numerous fusion methods often prioritize visual quality above semantic information. To address this problem, we present a Semantic Feature Interactive Learning Network (SFINet) for full-time infrared and visible images. The SFINet encompasses an image fusion network and an image segmentation network through a Semantic Feature Interaction (SFI) module. The image fusion network employs Multi-scale Feature Extraction (MFE) modules to capture global and local information at multiple scales. Meanwhile, it performs an adaptive fusion of complementary information using a Dual Attention Feature Fusion (DAFF) module. The image segmentation network guides the image fusion network using the SFI module for semantic feature interaction. Comparative results prove that the proposed method is superior to state-of-the-art (SOTA) models in image fusion and semantic segmentation tasks. The code is available at https://github.com/ songwenhao123/SFINet.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001337845400001  
Indexed Date:  
2025-01-23  
Institution Address:  
Shandong Univ Technol, Sch Elect & Elect Engn, Zibo 255000, Shandong, Peoples R China  
Institution Address:  
Queen Mary Univ London, Sch Elect Engn & Comp Sci, London E1 4NS, England  
Institution Address:  
Royal Mil Coll Canada, Dept Math & Comp Sci, Kingston, ON K7K 7B4, Canada  
Institution Address:  
Incheon Natl Univ, Dept Embedded Syst Engn, Incheon 22012, South Korea  
Institution:  
Institution:  
Institution:  
Institution:

### 705. M<SUP> 2</SUP>RNet: Multi-modal and multi-scale refined network for

摘要: Salient object detection is a fundamental topic in computer vision, which has promising application prospects. The previous methods based on RGB-D may potentially suffer from the incompatibility of multi-modal feature fusion and the insufficiency of multi-scale feature aggregation. To tackle these two dilemmas, we propose a novel multi-modal and multi-scale refined network (M 2 RNet). Specifically, three essential components are presented in this network. The nested dual attention module (NDAM) explicitly exploits the combined features of RGB and depth flows. The adjacent interactive aggregation module (AIAM) gradually integrates the neighbor features of high, middle and low levels. The joint hybrid optimization loss (JHOL) makes the predictions have a prominent outline. Extensive experiments quantitatively and qualitatively demonstrate that our method outperforms other state-of-the-art approaches.(c) 2022 Elsevier Ltd. All rights reserved.  
Web of Science 核心合集中的 "被引频次":  
20  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
20  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:000891815700002  
Indexed Date:  
2022-12-25  
Institution Address:  
Zhejiang Sci Tech Univ, Sch Comp Sci & Technol, Hangzhou 310018, Peoples R China  
Institution Address:  
Nankai Univ, Coll Comp Sci, Tianjin 300350, Peoples R China  
Institution Address:  
Nankai Univ, Coll Artificial Intelligence, Tianjin 300350, Peoples R China  
Institution:  
Institution:  
Institution:

### 706. A diffusion model multi-scale feature fusion network for imbalanced

摘要: Background and objective: Medicine image classification are important methods of traditional medical image analysis, but the trainable data in medical image classification is highly imbalanced and the accuracy of medical image classification models is low. In view of the above two common problems in medical image classification. This study aims to: (i) effectively solve the problem of poor training effect caused by the imbalance of class imbalanced data sets. (ii) propose a network framework suitable for improving medical image classification results, which needs to be superior to existing methods. Methods: In this paper, we put in the diffusion model multi-scale feature fusion network (DMSFF), which mainly uses the diffusion generation model to overcome imbalanced classes (DMOIC) on highly imbalanced medical image datasets. At the same time, it is processed according to the cropped image augmentation strategy through cropping (IASTC). Based on this, we use the new dataset to design a multi-scale feature fusion network (MSFF) that can fully utilize multiple hierarchical features. The DMSFF network can effectively solve the problems of small and imbalanced samples and low accuracy in medical image classification. Results: We evaluated the performance of the DMSFF network on highly imbalanced medical image classification datasets APTOS2019 and ISIC2018. Compared with other classification models, our proposed DMSFF network achieved significant improvements in classification accuracy and F1 score on two datasets, reaching 0.872, 0.731, and 0.906, 0.836, respectively. Conclusions: Our newly proposed DMSFF architecture outperforms existing methods on two datasets, and verifies the effectiveness of generative model inverse balance for imbalance class datasets and feature enhancement by multi-scale feature fusion. Further, the method can be applied to other class imbalanced data sets where the results will be improved.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0169-2607  
eISSN:  
1872-7565  
入藏号:  
WOS:001303138800001  
PubMed ID:  
39205335  
Indexed Date:  
2024-09-07  
Institution Address:  
Guangxi Univ, Sch Comp Elect & Informat, Nanning 530004, Guangxi, Peoples R China  
Institution Address:  
Univ Oulu, Ctr Machine Vis & Signal Anal, FI-90014 Oulu, Finland  
Institution Address:  
Guangxi Acad Sci, Big Data & Intelligent Comp Res Ctr, Nanning 530007, Peoples R China  
Institution Address:  
Nanning Normal Univ, Sch Comp & Informat Engn, Nanning 530299, Peoples R China  
Institution Address:  
Guangxi Univ, Guangxi Key Lab Multimedia Commun Network Technol, Nanning 530004, Guangxi, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 707. Prohibited Item Detection Within X-Ray Security Inspection Images Based

摘要: Automatic detection of prohibited items is vital in helping security staff be more efficient while improving the public safety index. However, prohibited item detection within X-ray security inspection images is limited by various factors, including the imbalance distribution of categories, diversity of prohibited item scales, and overlap between items. In this paper, we propose to leverage the Poisson blending algorithm with the Canny edge operator to alleviate the imbalance distribution of categories maximally in the X-ray images dataset. Based on this, we improve the cascade network to deal with the other two difficulties. To address the prohibited scale diversity problem, we propose the Re-BiFPN feature fusion method, which includes a coordinate attention atrous spatial pyramid pooling (CA-ASPP) module and a recursive connection. The CA-ASPP module can implicitly extract direction-aware and position-aware information from the feature map. The recursive connection feeds the CA-ASPP module processed multi-scale feature map to the bottom-up backbone layer for further multi-scale feature extraction. In addition, a Rep-CIoU loss function is designed to address the overlapping problem in X-ray images. Extensive experimental results demonstrate that our method can successfully identify ten types of prohibited items, such as Knives, Scissors, Pressure, etc. and achieves 83.4 degrees 1o of mAP, which is 3.8 degrees 1o superior to the original cascade network. Moreover, our method outperforms other mainstream methods by a significant margin.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0916-8508  
eISSN:  
1745-1337  
入藏号:  
WOS:001251848300016  
Indexed Date:  
2024-06-29  
Institution Address:  
Yamaguchi Univ, Grad Sch East Asian Studies, Yamaguchi 7538514, Japan  
Institution Address:  
Zhejiang Sci Tech Univ, Phys Dept, Hangzhou 310018, Peoples R China  
Institution Address:  
Shunan Univ, Fac Informat Sci, 745 8566 Japan, 745-8566, Shunan, Japan  
Institution Address:  
Yamaguchi Univ, Fac Educ, 753 8513 Japan, 753-8513, Yamaguchi, Japan  
Institution Address:  
Yamaguchi Univ, Yamaguchi 7538511, Japan  
Institution:  
Institution:  
Institution:  
Shunan Univ  
Institution:  
Institution:

### 708. Lightweight target detection for the field flat jujube based on improved

摘要: The efficient detection of the flat jujube in a complex natural environment has great significance in intelligent agricultural operations. Aiming at the problems of the low detection efficiency of field flat jujubes and complex target detection algorithms that are difficult to deploy on low-cost equipment, an improved lightweight algorithm based on You Only Look Once (YOLOv5) is proposed. First, the method screens for the multiscale detection structure that is suitable for the flat jujube by adjusting the number of layers of target detection, which improves the accuracy of detection and reduces the nuisance parameter. Then, multiscale feature fusion is achieved more efficiently by using the bidirectional feature pyramid network (BiFPN), and the feature extraction capability of the model is further improved by introducing a dual coordinate attention mechanism. Finally, the method reduces the difficulties of the model by introducing depthwise separable convolution and adding a ghost module after upsampling layers. The experimental results showed that the mean average precision (mAP) and model size of the lightweight network reached 97.2 % and 7.1 MB. Compared with the YOLOv5 baseline network, the parameters decreased by 49.15 %, while the mAP increased by 1.8 %. The method further improved algorithm performance and reduced computational cost compared with the mainstream one-stage target detection algorithms of the YOLOv5s, YOLOx\_s, YOLOv4, YOLOv3 and single shot multibox detector (SSD). Compared to these algorithms, the mAP of the proposed improved model increased by 1.8 %, 0.9 %, 5.5 %, 6.5 % and 2.9 %, respectively. Meanwhile, the model size was compressed by 49.15 %, 73.99 %, 94.42 %, 94.24 % and 86.69 %, respectively. The improved algorithm has higher detection accuracy, while reducing the calculations and parameters, which reduces the dependence on hardware and provides a reference for deploying automated picking of the field flat jujube.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
50  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
4  
在 BIOSIS Citation Index 中的被引频次:  
9  
被引频次合计:  
54  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:000888854400003  
Indexed Date:  
2022-12-21  
Institution Address:  
Shanxi Agr Univ, Coll Agr Engn, Taigu 030800, Peoples R China  
Institution:

### 709. SCB-YOLOv5: a lightweight intelligent detection model for athletes'

摘要: Intelligent detection of athlete behavior is beneficial for guiding sports instruction. Existing mature target detection algorithms provide significant support for this task. However, large-scale target detection algorithms often encounter more challenges in practical application scenarios. We propose SCB-YOLOv5, to detect standardized movements of gymnasts. First, the movements of aerobics athletes were captured, labeled using the labelImg software, and utilized to establish the athlete normative behavior dataset, which was then enhanced by the dataset augmentation using Mosaic9. Then, we improved the YOLOv5 by (1) incorporating the structures of ShuffleNet V2 and convolutional block attention module to reconstruct the Backbone, effectively reducing the parameter size while maintaining network feature extraction capability; (2) adding a weighted bidirectional feature pyramid network into the multiscale feature fusion, to acquire precise channel and positional information through the global receptive field of feature maps. Finally, SCB-YOLOv5 was lighter by 56.9% than YOLOv5. The detection precision is 93.7%, with a recall of 99% and mAP value of 94.23%. This represents a 3.53% improvement compared to the original algorithm. Extensive experiments have verified that our method. SCB-YOLOv5 can meet the requirements for on-site athlete action detection. Our code and models are available at https://github.com/qingDu1/SCB-YOLOv5.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2045-2322  
入藏号:  
WOS:001202213000010  
PubMed ID:  
38616199  
Indexed Date:  
2024-05-02  
Institution Address:  
Univ South China, Sch Resource Environm & Safety Engn, Hengyang 421001, Peoples R China  
Institution Address:  
Hunan Inst Engn, Sch Sports Sci & Engn, Xiangtan 411100, Peoples R China  
Institution Address:  
Hunan Inst Engn, Sch Elect Informat Engn, Xiangtan 411100, Peoples R China  
Institution:  
Institution:  
Institution:

### 710. RCFusion: Fusing 4-D Radar and Camera With Bird's-Eye View Features for

摘要: Camera and millimeter-wave (MMW) radar fusion is essential for accurate and robust autonomous driving systems. With the advancement of radar technology, next-generation high-resolution automotive radar, i.e., 4-D radar, has emerged. In addition to the target range, azimuth, and Doppler velocity measurements of traditional radar, 4-D radar provides elevation measurement to create a denser "point cloud." In this study, we propose a camera and 4-D radar fusion network called RCFusion, which achieves multimodal feature fusion under a unified bird's-eye view (BEV) space to accomplish 3-D object detection tasks. In the camera stream, multiscale feature maps are obtained by the image backbone and feature pyramid network (FPN); they are then converted into orthographic feature maps by an orthographic feature transform (OFT). Next, enhanced and fine-grained image BEV features are obtained via a designed shared attention encoder. Meanwhile, in the 4-D radar stream, a newly designed component named radar PillarNet efficiently encodes the radar features to generate radar pseudo-images, which are fed into the point cloud backbone to create radar BEV features. An interactive attention module (IAM) is proposed for the fusion stage, which outputs a valid fusion of the two-modal BEV features. Finally, a generic detection head predicts the object classes and locations. The proposed RCFusion is validated on the TJ4DRadSet and view-of-delft (VoD) datasets. The experimental results and analysis show that the proposed method can effectively fuse camera and 4-D radar features to achieve robust detection performance.  
Web of Science 核心合集中的 "被引频次":  
32  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
32  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001005752200027  
Indexed Date:  
2023-07-18  
Institution Address:  
Tongji Univ, Sch Automot Studies, Shanghai 201804, Peoples R China  
Institution Address:  
Zhejiang Univ City Coll, Sch Informat & Elect, Hangzhou 310015, Peoples R China  
Institution:  
Institution:

### 711. HAFREE: A Heatmap-Based Anchor-Free Detector for Apple Defect Detection

摘要: Accurate inspection of subtle defects on apple surfaces is necessary in agricultural engineering. However, existing methods often rely on expensive equipment and encounter difficulty in detecting small defect areas effectively. To address this challenge, we introduce the Subtle Surface Defects in Apples (SSDA) dataset, a custom dataset specifically collected and annotated for two defect types: scratches and pest damage. For benchmarking, we propose a heatmap-based anchor-free (HAFREE) detector, a novel end-to-end object detection architecture designed to localize subtle defects on apple surfaces. Unlike prior methods that rely on anchor boxes, HAFREE employs a heatmap-based approach to represent defects as keypoints using two-dimensional Gaussian heatmaps. We introduce a multiscale feature fusion block, targeting small objects by incorporating local and global contextual information. To mitigate overfitting, we also implement a patch training strategy incorporating full images and cropped patches during training as a regularizer. The proposed method achieves a mAP50 of 50.05% on the SSDA dataset, outperforming one- and two-stage anchor-based detectors and previous anchor-free approaches. Code is available at: https://github.com/nbngochan/HAFREE.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001375790900022  
Indexed Date:  
2025-02-26  
Institution Address:  
Kwangwoon Univ, Dept Elect Convergence Engn, Seoul 01897, South Korea  
Institution Address:  
Chonnam Natl Univ, Dept Intelligent Elect & Comp Engn, Gwangju 61186, South Korea  
Institution Address:  
AISeed Inc, Res Ctr, Gwangju 61186, South Korea  
Institution Address:  
Air Univ Islamabad, Fac Comp & AI, Dept Creat Technol, Islamabad 44000, Pakistan  
Institution:  
Institution:  
Institution:  
AISeed Inc  
Institution:

### 712. STGWN: Enhanced spatiotemporal wave forecasting using multiscale

摘要: This paper addresses the challenge of low accuracy in wave height prediction due to inadequate consideration of irregular topological structures and spatio-temporal dependencies of wave measurement points. We propose a novel spatio-temporal wave prediction method, named the Spatio-Temporal Graph Wave Network (STGWN). This model employs graph convolution instead of conventional 2D convolution, enabling spatial domain modeling of wave fields on a graph structure. Simultaneously, it utilizes LSTM for temporal modeling, effectively enhancing the representation of complex dynamic spatio-temporal correlations among wave nodes. Additionally, a dynamic wave position-aware mechanism is constructed to better identify feature variations among different wave nodes. Furthermore, a multi-scale feature fusion mechanism is designed within the model to prevent the risk of crucial temporal feature disappearance during the extraction of spatial dimension features. The proposed spatio-temporal method for Significant Wave Height (SWH) prediction is validated on a wave dataset comprising four different node distribution characteristics within the study area. Results demonstrate that the STGWN model outperforms comparative models in terms of prediction accuracy and stability across multiple evaluation metrics in the four experimental setups.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
0141-1187  
eISSN:  
1879-1549  
入藏号:  
WOS:001188184300001  
Indexed Date:  
2024-03-31  
Institution Address:  
Dalian Univ Technol, Sch Mech & Aerosp Engn, Dept Engn Mech, State Key Lab Struct Anal Optimizat & CAE Software, Dalian 116023, Liaoning, Peoples R China  
Institution Address:  
Dalian Univ Technol, Ningbo Inst, Ningbo 315000, Peoples R China  
Institution:  
Institution:

### 713. Grasp Pose Estimation Based on Multi-Scale Feature Fusion

摘要: In order to improve the success rate of robot grasping,a multi-scale feature fusion method for robot grasping pose estimation is proposed.The method takes RGD information as input,uses ResNet-50 backbone network and integrates FPN(feature pyramid networks)to obtain multi- scale features as the input of grasping generation network to generate grasping candidate frame.The grasping direction coordinates are mapped as the classification task of grasping direction,and ROI Align is used to extract the region of interest,evaluate the grasping candidate box,and obtain the optimal grasping pose of the target.In order to verify the effectiveness of the proposed algorithm,the pose estimation experiment based on Cornell data set is carried out,and the accuracy of pose estimation reaches 96.9%.Based on the Inter RealSense D415 depth camera and UR5 manipulator,a real object platform is built.In the real scene,multiple grasping experiments are carried out on the diverse objects randomly placed in the real scene.The results show that the detection success rate of grasping target is 95.8%,and the success rate of robot grasping is 90.2%.  
摘要:  
抓取目标多样性、位姿随机性严重制约了机器人抓取的任务适应性,为提高机器人抓取成功率,提出一种融合多尺度特征的机器人抓取位姿估计方法。该方法以RGD信息为输入,采用ResNet-50主干网络,融合FPN(feature pyramid networks)获得多尺度特征作为抓取生成网络的输入,以生成抓取候选框;并将抓取方向坐标映射为抓取方向的分类任务,使用ROI Align进行感兴趣区域提取,评估抓取候选框,获取目标的最优抓取位姿。为验证算法有效性,基于康奈尔抓取数据集开展了抓取位姿估计实验,仿真抓取位姿估计准确度达到96.9%。基于Inter RealSense D415深度相机和UR5机械臂搭建了实物平台,在真实场景下对位姿随机摆放的多样性目标物体进行多次抓取实验,结果显示抓取目标检测成功率为95.8%,机器人抓取成功率为90.2%。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1002-8331  
入藏号:  
CSCD:7219720  
Indexed Date:  
2022-08-05  
Institution Address:  
西南科技大学信息工程学院, 绵阳, 四川 621010, 中国  
Institution Address:  
西南科技大学信息工程学院;;中国科学技术大学信息科学技术学院, ;;, 绵阳;;合肥, 四川;; 621010;;230026, 中国  
Institution Address:  
School of Information Engineering,Southwest University of Science and Technology, Mianyang, Sichuan 621010, China  
Institution Address:  
School of Information Engineering,Southwest University of Science and Technology;;School of Information Science and Technology,University of Science and Technology of China, ;;, Mianyang;;Hefei, Sichuan;; 621010;;230026  
Institution:  
西南科技大学信息工程学院  
Institution:  
西南科技大学信息工程学院;;中国科学技术大学信息科学技术学院  
Institution:  
School of Information Engineering,Southwest University of Science and Technology  
Institution:  
School of Information Engineering,Southwest University of Science and Technology;;School of Information Science and Technology,University of Science and Technology of China

### 714. Lightweight Human Pose Estimation with Cascaded Channel Attention

摘要: Aiming at the problem of serious loss of accuracy in the lightweighting process of the current human pose estimation model,a lightweight human pose estimation model that introduces cascaded channel attention is proposed using the high resolution network (HRNet) as a baseline.Firstly,a cascading channel attention that maintains internal high-resolution features is constructed so as to learn the importance of each channel of the input features to improve the model representation.Secondly,the residual module with high arithmetic complexity in HRNet stages 2,3,and 4 is replaced by designing a lightweight deepwise convolutional transform module based on the structure of the MetaFormer.Furthermore,a multi-scale feature fusion method is designed to reduce the loss of semantic information of multi-dimensional features in the original fusion method of HRNet.Finally,unbiased data processing is used to eliminate offset errors caused by the process of encoding the heat map at key points.Experimental results from the COCO\_2017 validation set show that the proposed model reduces the number of model parameters and floating-point operations by 90.2% and 83.1%,respectively,compared to the benchmark model with a 2 percentage points decrease in AP,and achieves the optimal accuracy among the lightweight models with an AP of 71.4%.  
摘要:  
针对当前人体姿态估计模型在轻量化过程中精度损失严重的问题,以高分辨率网络(HRNet)为基线提出一种引入级联通道注意力的轻量化人体姿态估计模型。构建一种保持内部高分辨率特征的级联通道注意力,学习输入特征各通道的重要性来提高模型表征能力;通过设计一种基于MetaFormer结构的轻量级深度卷积变换模块来替换HRNet阶段2、3、4中运算复杂度较高的残差模块;设计一种多尺度特征融合方法减少HRNet原融合方法中的多维特征语义信息损失;采用无偏数据处理来消除关键点热力图编码过程中导致的偏移误差。COCO\_2017验证集的实验结果表明,所提出的模型同基准模型相比,在AP降低2个百分点的情况下,模型参数量和浮点运算量分别减少了90.2%和83.1%,并且以AP为71.4%的表现在轻量化模型中达到精度最优。  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1002-8331  
入藏号:  
CSCD:7794561  
Indexed Date:  
2024-11-01  
Institution Address:  
西安工业大学电子信息工程学院, 西安, 陕西 710021, 中国  
Institution Address:  
西安工业大学发展规划处, 西安, 陕西 710021, 中国  
Institution Address:  
School of Electronics and Information Engineering,Xi'an Technological University, Xi'an, Shaanxi 710021, China  
Institution Address:  
Development Planning Department,Xi'an Technological University, Xi'an, Shaanxi 710021, China  
Institution:  
西安工业大学电子信息工程学院  
Institution:  
西安工业大学发展规划处  
Institution:  
School of Electronics and Information Engineering,Xi'an Technological University  
Institution:  
Development Planning Department,Xi'an Technological University

### 715. A shared multi-scale lightweight convolution generative network for

摘要: Time series forecasting is an important time series data mining technique. Among them, multivariate time series (MTS) forecasting has received extensive attention in many fields. However, many existing MTS forecasting models usually rely on a large amount of labeled data for model training, and data collection and labeling are difficult in real systems. The insufficient amount of data makes it difficult for the model to fully learn the intrinsic patterns and features of the data, which not only increases the prediction error, but also makes it hard to obtain satisfactory prediction results. To address this challenge, we propose a shared multi-scale lightweight convolution generative (SMLCG) network for few-shot multivariate time series forecasting by using samples generation strategy. The overall goal is to design a shared multi-scale feature generation prediction framework that generates data highly similar to the original sample and enriches the training sample to improve prediction accuracy. Specifically, the MTS is divided into different scales, and the multi-scale feature fusion module is utilized to capture and fuse the MTS information indifferent spatial dimensions to eliminate the heterogeneity among the data. Then, the key information in the multi-scale features is captured by a lightweight convolution generative network, and the feature weights are dynamically assigned to explore the change information. In addition, a spatio-temporal memory module is designed based on the parameter sharing strategy to capture the spatio-temporal dynamic relationship of sequences by learning the common knowledge in multi-scale features, thus improving the robustness and generalization ability. Through comprehensive experiments on four publicly available datasets and comparisons with other reported models, it is demonstrated that the SMLCG model can efficiently generate approximate samples in the few-shot case and provide excellent prediction results. The architecture of SMLCG serves as a valuable reference for practical solutions to address the few-shot problem in multivariate time series.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1568-4946  
eISSN:  
1872-9681  
入藏号:  
WOS:001353650100001  
Indexed Date:  
2024-11-19  
Institution Address:  
Southwest Jiaotong Univ, Sch Comp & Artificial Intelligence, Chengdu 610031, Sichuan, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, Sch Informat Sci & Technol, Chengdu 610031, Sichuan, Peoples R China  
Institution Address:  
Mfg Ind Chains Collaborat & Informat Support Techn, Chengdu 610031, Sichuan, Peoples R China  
Institution:  
Institution:  
Institution:  
Mfg Ind Chains Collaborat & Informat Support Techn

### 716. A Real-Time Negative Obstacle Detection Method for Autonomous Trucks in

摘要: Negative obstacles such as potholes and road collapses on unstructured roads in open-pit mining areas seriously affect the safe transportation of autonomous trucks. In this paper, we propose a real-time negative obstacle detection method for self-driving trucks in open-pit mines. By analyzing the characteristics of road negative obstacles in open-pit mines, a real-time target detection model based on the Yolov4 network was built. It uses RepVGG as the backbone feature extraction network, applying SimAM space and a channel attention mechanism to negative obstacle multiscale feature fusion. In addition, the classification and prediction modules of the network are optimized to improve the accuracy with which it detects negative obstacle targets. A non-maximum suppression optimization algorithm (CIoU Soft Non-Maximum Suppression, CS-NMS) is proposed in the post-processing stage of negative obstacle detection. The CS-NMS calculates the confidence of each detection frame with weighted optimization to solve the problems of encountering obscure negative obstacles or poor positioning accuracy of the detection boxes. The experimental results show that this research method achieves 96.35% mAP for detecting negative obstacles on mining roads with a real-time detection speed of 69.3 fps, and that it can effectively identify negative obstacles on unstructured roads in open-pit mines with complex backgrounds.  
Web of Science 核心合集中的 "被引频次":  
9  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
eISSN:  
2071-1050  
入藏号:  
WOS:000908998500001  
Indexed Date:  
2023-01-23  
Institution Address:  
Xian Univ Architecture & Technol, Sch Resource Engn, Xian 710055, Peoples R China  
Institution Address:  
Xian Univ Architecture & Technol, Xian Key Lab Intelligent Ind Percept Comp & Decis, Xian 710055, Peoples R China  
Institution:  
Institution:

### 717. PTCDet: advanced UAV imagery target detection

摘要: Object detection in drone aerial images is challenging due to the difficulty of small object detection and complex backgrounds. To address these issues, this paper proposes an improved object detection model, perception and target capture detector (PTCDet), to increase detection accuracy and robustness in complex scenes. Specifically, the proposed multiple feature extraction attention (MFEA) module significantly enhances the ability of the model to detect small objects through multidimensional feature map augmentation. The weighted perceptive field augmentation (WPFA) module is designed to improve the contextual awareness and feature representation of the model, optimizing detection accuracy for small objects. Based on the multiscale feature fusion structure, an enhanced scale fusion detection (ESFD) module is used to improve small object detection by generating larger scale feature maps. Ultimately, the inner focaler IoU loss (INFL) function effectively accelerates the regression of detection bounding boxes, enhancing the generalization ability and overall detection performance of the model. The experimental results on three public datasets demonstrate that PTCDet outperforms other detection algorithms. For example, on the VisDrone dataset, compared with the baseline model YOLOv8, map@0.5 and map@0.5:0.95 are improved by 6.21% and 4.21%, respectively. PTCDet exhibits excellent performance in addressing complex backgrounds and small object detection, providing an effective and robust solution for object detection tasks in drone aerial images.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2045-2322  
入藏号:  
WOS:001352487800105  
PubMed ID:  
39521873  
Indexed Date:  
2024-11-23  
Institution Address:  
Hebei Univ Sci & Technol, Coll Informat Sci & Engn, Shijiazhuang 050018, Peoples R China  
Institution:

### 718. COD-YOLO: An Efficient YOLO-Based Detector for Laser Chip Catastrophic

摘要: High-power semiconductor lasers play a crucial role in optical communication systems, and their reliability is key to the normal operation of the system. Catastrophic Optical Damage generated during operation is a major factor affecting chip performance and lifetime. Accurate detection of the location and development process of damage, along with the study of failure mechanisms and degradation modes, is a pressing issue. We propose an intelligent analysis approach based on the YOLO architecture for defect detection in laser chip Catastrophic Optical Damage, named COD-YOLO. To overcome challenges such as the similarity of defect features, complexity of background features, and inaccurate spatial positioning, the network employs deformable convolutional and channel attention. This adaptive approach captures rich feature representations and simultaneously addresses long-distance dependencies and adaptive spatial aggregation. Combining spatial-content-based upsampling in model neck achieves multiscale feature fusion, improving perception and understanding through the integration of semantic and positional information. Furthermore, due to the lack of fine-grained information, IoU metrics are highly sensitive to the positional deviation of tiny defects. Combining tiny object detection loss function to measure the regression of bounding boxes, adapting to variations in defect scales, experimental findings demonstrate that COD-YOLO outperforms other competing methods in detecting Catastrophic Optical Damage in the active region of the laser chip.  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2193-567X  
eISSN:  
2191-4281  
入藏号:  
WOS:001281860500001  
Indexed Date:  
2024-08-06  
Institution Address:  
Taiyuan Univ Technol, Coll Elect Informat & Opt Engn, Taiyuan 030024, Peoples R China  
Institution Address:  
Taiyuan Univ Technol, Coll Comp Sci & Technol, Coll Data Sci, Taiyuan 030024, Peoples R China  
Institution Address:  
Taiyuan Univ Technol, Coll Elect & Power Engn, Taiyuan 030024, Peoples R China  
Institution Address:  
Key Lab Big Data Fus Anal & Applicat Shanxi Prov, Taiyuan 030024, Peoples R China  
Institution Address:  
Intelligent Percept Engn Technol Ctr Shanxi, Taiyuan 030024, Peoples R China  
Institution Address:  
Accelink Technol Co Ltd, Wuhan 430000, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Key Lab Big Data Fus Anal & Applicat Shanxi Prov  
Institution:  
Intelligent Percept Engn Technol Ctr Shanxi  
Institution:  
Accelink Technol Co Ltd

### 719. Multiscale Low-Light Image Enhancement Algorithm with Brightness

摘要: To address issues such as detail loss, artifacts, and unnatural appearance associated with current low-illumination image enhancement algorithms, a multiscale low-illumination image enhancement algorithm based on brightness equalization and edge enhancement is proposed in this study. Initially, an improved Sobel operator is employed to extract edge details, yielding an image with enhanced edge details. Subsequently, the brightness component (V) of the HSV color space is enhanced using Retinex, and brightness equalization is accomplished via improved Gamma correction, yielding an image with balanced brightness. The Laplacian weight graph, significance weight graph, and saturation weight graph are computed for the edge detail-enhanced image and brightness-balanced image, culminating in the generation of a normalized weight graph. This graph is then decomposed into a Gaussian pyramid, while the edge detail-enhanced image and brightness-balanced image are decomposed into a Laplacian pyramid. Finally, a multiscale pyramid fusion strategy is employed to merge the images, resulting in the final enhanced image. Experimental results demonstrate that the proposed algorithm outperforms existing algorithms on the LOL dataset in terms of average peak signal to noise ratio, structural similarity, and naturalness image quality evaluator. This algorithm effectively enhances the contrast and clarity of low-illumination images, resulting in images with richer detail information, improved color saturation, and considerably enhanced quality.  
摘要:  
为解决现有低照度图像增强算法存在的细节信息丢失、伪影和自然度较差等问题,提出一种亮度均衡与边缘强化的多尺度低照度图像增强算法。首先,利用改进的Sobel算子提取边缘细节,得到边缘细节增强图像;然后,对图像HSV颜色空间亮度V进行Retinex增强,并用改进的Gamma校正进行亮度均衡化处理,得到亮度均衡图像;接下来,计算边缘细节增强图像和亮度均衡图像的拉普拉斯权重图、显著性权重图和饱和度权重图,进而得到归一化权重图;最后,将归一化权重图分解成高斯金字塔,将边缘细节增强图像和亮度均衡图像分解成拉普拉斯金字塔,采用多尺度金字塔融合策略进行图像融合,得到最终增强图像。实验结果表明,所提算法在LOL数据集上的峰值信噪比、结构相似性和自然度图像质量评估器的平均值均优于其他算法,且能够有效提升低照度图像的对比度、清晰度,增强后的图像细节信息更丰富、色彩饱和度更好,图像质量提升明显。  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7783066  
Indexed Date:  
2024-10-18  
Institution Address:  
辽宁工程技术大学软件学院;;辽宁工程技术大学基础教学部, ;;, 葫芦岛;;葫芦岛, 辽宁;;辽宁 125105;;125105, 中国  
Institution Address:  
辽宁工程技术大学软件学院, 葫芦岛, 辽宁 125105, 中国  
Institution Address:  
中煤科工集团沈阳研究院有限公司;;煤矿安全技术国家重点实验室, ;;煤矿安全技术国家重点实验室, 抚顺;;抚顺, 辽宁;;辽宁 113122;;113122, 中国  
Institution Address:  
School of Software,Liaoning Technical University;;Department of Basic Education,Liaoning Technical University, ;;, Huludao;;Huludao, Liaoning;;Liaoning 125105;;125105  
Institution Address:  
School of Software,Liaoning Technical University, Huludao, Liaoning 125105, China  
Institution Address:  
China Coal Technology & Engineering Group Shenyang Research Institute;;State Key Laboratory of Coal Mine Safety Technology, ;;State Key Laboratory of Coal Mine Safety Technology, Fushun;;Fushun, Liaoning;;Liaoning 113122;;113122  
Institution:  
辽宁工程技术大学软件学院;;辽宁工程技术大学基础教学部  
Institution:  
辽宁工程技术大学软件学院  
Institution:  
中煤科工集团沈阳研究院有限公司;;煤矿安全技术国家重点实验室  
Institution:  
School of Software,Liaoning Technical University;;Department of Basic Education,Liaoning Technical University  
Institution:  
School of Software,Liaoning Technical University  
Institution:  
China Coal Technology & Engineering Group Shenyang Research Institute;;State Key Laboratory of Coal Mine Safety Technology

### 720. S3INet: Semantic-Information Space Sharing Interaction Network for

摘要: The detecting arbitrary shape text is a challenging task due to the significant variation in text shape, size, and aspect ratio, as well as the complexity of scene backgrounds. The enhancing feature extraction capabilities is essential for the boosting text detection accuracy. However, traditional text feature extraction methods face several issues, including insufficient multiscale feature fusion, limited information transfer between different feature levels, and constrained receptive field expansion when using asymmetric convolutional kernels for long text detection. To address these challenges, this article introduces an arbitrarily shaped scene text detector called the semantic-information space sharing interaction network (S3INet). The proposed network leverages the semantic-information space sharing module (S3M) to generate a single-level feature map capable of capturing multiscale features with rich semantic information and prominent foreground elements. In addition, we propose the multibranch parallel asymmetric convolutional module (MPACM) group to enhance the representation of text features, thereby further enhancing text detection performance. Extensive experimental evaluations on five publicly available natural scene text datasets (CTW-1500, Total-Text, MSRA-TD500, ICDAR2015, and ICDAR2017-MLT) and two traffic text datasets (CTST-1600 and TPD) demonstrate the superiority of our method. The results indicate that S3INet significantly outperforms most existing state-of-the-art methods in both accuracy and robustness. The code will be released at: https://github.com/runminwang/S3INet.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2162-2388  
入藏号:  
MEDLINE:40036453  
PubMed ID:  
40036453  
Indexed Date:  
2025-03-07

### 721. Task-Adaptive Embedding Learning with Dynamic Kernel Fusion for Few-Shot

摘要: The central goal of few-shot scene classification is to learn a model that can generalize well to a novel scene category (UNSEEN) from only one or a few labeled examples. Recent works in the Remote Sensing (RS) community tackle this challenge by developing algorithms in a meta-learning manner. However, most prior approaches have either focused on rapidly optimizing a meta-learner or finding good similarity metrics while overlooking the embedding power. Here we propose a novel Task-Adaptive Embedding Learning (TAEL) framework that complements the existing methods by giving full play to feature embedding's dual roles in few-shot scene classification-representing images and constructing classifiers in the embedding space. First, we design a Dynamic Kernel Fusion Network (DKF-Net) that enriches the diversity and expressive capacity of embeddings by dynamically fusing information from multiple kernels. Second, we present a task-adaptive strategy that helps to generate more discriminative representations by transforming the universal embeddings into task-adaptive embeddings via a self-attention mechanism. We evaluate our model in the standard few-shot learning setting on two challenging datasets: NWPU-RESISC4 and RSD46-WHU. Experimental results demonstrate that, on all tasks, our method achieves state-of-the-art performance by a significant margin.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
17  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
17  
eISSN:  
2072-4292  
入藏号:  
WOS:000719117800001  
Indexed Date:  
2021-11-24  
Institution Address:  
Northwestern Polytech Univ, Sch Comp Sci, Natl Engn Lab Integrated Aerosp Ground Ocean Big, Shaanxi Prov Key Lab Speech & Image Informat Proc, Xian 710129, Peoples R China  
Institution Address:  
Oklahoma State Univ, Sch Elect & Comp Engn, Stillwater, OK 74078 USA  
Institution Address:  
Yanan Univ, Sch Phys & Elect Informat, Yanan 716000, Peoples R China  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Commun & Informat Engn, Xian 710121, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 722. A feature pyramid network with adaptive fusion strategy and enhanced

摘要: In order to better detect objects of different scales, detectors need different resolutions and inputs from different receptive fields. Currently, advanced detectors usually combine the structure of feature pyramid to achieve the fusion of multi-scale object features. Top-down and bottom-up network structure is the basic strategy of multi-scale feature extraction. Although the feature pyramid network(FPN) can alleviate the contradiction between resolution and receptive field to a certain extent, the existing models based on FPN tend to ignore the contradictory information between different layers in the fusion process, and some fuzzy boundary information is also prone to lose features in top-down propagation. This paper first introduces the detector, then analyzes the defects behind the feature pyramid network, and finally proposes a feature pyramid network(SG-FPN) with adaptive fusion strategy and enhanced semantic information to solve these problems. The validity of our model is verified on mainstream data sets, and the performance is superior compared with other state-of-the-art methods.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0942-4962  
eISSN:  
1432-1882  
入藏号:  
WOS:001243657800001  
Indexed Date:  
2024-06-17  
Institution Address:  
Tianjin Univ Technol, Tianjin Key Lab Intelligence Comp & Novel Software, 391 West Binshui Rd, Tianjin 300384, Peoples R China  
Institution:

### 723. Efficient algorithm for directed text detection based on rotation

摘要: A more effective directed text detection algorithm is proposed for the problem of low accuracy in detecting text with multiple sources, dense distribution, large aspect ratio and arbitrary alignment direction in the industrial intelligence process. The algorithm is based on the YOLOv5 model architecture, inspired by the idea of DenseNet dense connection, a parallel cross-scale feature fusion method is proposed to overcome the problem of blurring the underlying feature semantic information and deep location information caused by the sequential stacking approach and to improve the multiscale feature information extraction capability. Furthermore, a rotational decoupling border detection module, which decouples the rotational bounding box into horizontal bounding box during positive sample matching, is provided, overcoming the angular instability in the process of matching the rotational bounding box with the horizontal anchor to obtain higher-quality regression samples and improve the precision of directed text detection. The MSRA-TD500 and ICDAR2015 datasets are used to evaluate the method, and results show that the algorithm measured precision and F1-score of 89.2% and 88.1% on the MSRA-TD500 dataset, respectively, and accuracy and F1-score of 90.6% and 89.3% on the ICDAR2015 dataset, respectively. The proposed algorithm has better competitive ability than the SOTA text detection algorithm.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2376-5992  
入藏号:  
WOS:000996387100004  
PubMed ID:  
37346620  
Indexed Date:  
2023-06-15  
Institution Address:  
Fujian Univ Technol, Fujian Key Lab Intelligent Machining Technol & Equ, Fuzhou, Peoples R China  
Institution Address:  
Fujian Wuyi Leaf Tobacco Co Ltd, Shaowu, Peoples R China  
Institution:  
Institution:  
Fujian Wuyi Leaf Tobacco Co Ltd

### 724. Image super-resolution network based on multi-scale adaptive attention

摘要: Aiming at the problem that most image super-resolution methods cannot fully extract features by using single-scale convolution, an image super-resolution network based on multi-scale adaptive attention is proposed. To fully use the contextual information in each hierarchical feature, a multi-scale feature fusion block was designed, whose basic unit consists of an adaptive dual-scale block, a multi-path progressive interactive block, and an adaptive dual-dimensional attention sequentially in series. Firstly, the adaptive dual-scale block autonomously fused the features of two scales to obtain richer contextual features; secondly, the multi-path progressive interactive block interacted the output of the adaptive dual-scale block in a progressive way to improve the correlation between the contextual features; lastly, the adaptive dual-dimensional attention autonomously selected different dimensions of the attention to refine the output features, which makes the output features more discriminative. The experimental results show that on Set5, Set14, BSD100 and Urban100 test sets, the method of this paper improves the PSNR and SSIM quantitative metrics compared to other mainstream methods, especially for the Urban100 test set, where texture details are difficult to be recovered, the method of this paper improves PSNR and SSIM metrics by 0.05 dB and 0.004 5 respectively compared to the existing optimal method, SwinIR, with the scaling factor of \*4; in terms of visual effect, the reconstructed images in this paper have more texture details.  
摘要:  
针对大多数图像超分辨率重建方法利用单尺度卷积进行特征提取,导致特征提取不充分的问题,提出基于多尺度自适应注意力的图像超分辨率网络。为充分利用各个层次特征中的上下文信息,设计了多尺度特征融合块,其基本单元由自适应双尺度块、多路径渐进式交互块和自适应双维度注意力依次串联组成。首先,自适应双尺度块自主融合两个尺度的特征,获得了更丰富的上下文特征;其次,多路径渐进式交互块以渐进的方式交互自适应双尺度块的输出特征,提高了上下文特征之间的关联性;最后,自适应双维度注意力自主选择不同维度注意力细化输出特征后,提高了输出特征的鉴别力。实验结果表明,在Set5, Set14, BSD100和Urban100测试集上,本文方法在PSNR和SSIM定量指标上相比于其他主流方法相均有提升,尤其对于纹理细节难以恢复的Urban100测试集,本文方法在比例因子为\*4时,相较于现有最优方法SwinIR,PSNR和SSIM指标分别提升了0.05 dB和0.004 5;在视觉效果方面,本文的重建图像具有更多的纹理细节。  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
2  
ISSN:  
1004-924X  
入藏号:  
CSCD:7682888  
Indexed Date:  
2024-06-07  
Institution Address:  
河北工业大学人工智能与数据科学学院;;河北省控制工程技术研究中心, ;;河北省控制工程技术研究中心, ;;, 天津;;天津 300130;;300130, 中国  
Institution Address:  
河北工业大学人工智能与数据科学学院, 天津 300130, 中国  
Institution Address:  
School of Artificial Intelligence, Hebei University of Technology;;China Hebei Control Engineering Research Center, ;;China Hebei Control Engineering Research Center, ;;, Tianjin;;Tianjin 300130;;300130  
Institution Address:  
School of Artificial Intelligence, Hebei University of Technology, Tianjin 300130, China  
Institution:  
河北工业大学人工智能与数据科学学院;;河北省控制工程技术研究中心  
Institution:  
河北工业大学人工智能与数据科学学院  
Institution:  
School of Artificial Intelligence, Hebei University of Technology;;China Hebei Control Engineering Research Center  
Institution:  
School of Artificial Intelligence, Hebei University of Technology

### 725. Overcoming occlusions in complex environments to achieve robust

摘要: To effectively address the emotional well-being of elderly individuals living alone using of home care robotic systems, it is essential to possess the ability to precisely identify facial expressions within complex domestic settings. Facial expression recognition (FER) in complex environments faces significant challenges due to factors such as facial occlusions. To address this challenge, this paper proposes a method called the Dual-Branch Attention and Multi-Scale Feature Fusion Network (DAMFF-Net). First, we perform feature extraction on facial images and input the resulting feature maps into an improved dual-branch attention fusion module (DBAF) to capture long-range dependencies between different facial regions. Simultaneously, using the residual multi-scale module that we designed, we obtain fine-grained multi-scale features to ensure that both preceding and subsequent feature subsets contain rich scale information. Next, we globally fuse the feature maps from the feature extraction stage with those from the residual multi-scale module to enhance facial expression recognition accuracy in cases where certain facial feature regions are occluded. Finally, we employ decision-level fusion in order to deal with the categorization findings. Experiments were conducted on the RAF-DB, CK+ and AffectNet-7 datasets, and comparative results indicated that the proposed method achieved a respective enhancement of 5.79%, 6.68% and 5.86% in facial expression recognition accuracy.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2631-8695  
入藏号:  
WOS:001385259800001  
Indexed Date:  
2025-01-03  
Institution Address:  
Nanjing Inst Technol, Sch Automat, Nanjing 211167, Peoples R China  
Institution:

### 726. MSPV3D: Multi-Scale Point-Voxels 3D Object Detection Net

摘要: Autonomous vehicle technology is advancing, with 3D object detection based on point clouds being crucial. However, point clouds' irregularity, sparsity, and large data volume, coupled with irrelevant background points, hinder detection accuracy. We propose a two-stage multi-scale 3D object detection network. Firstly, considering that a large number of useless background points are usually generated by the ground during detection, we propose a new ground filtering algorithm to increase the proportion of foreground points and enhance the accuracy and efficiency of the two-stage detection. Secondly, given that different types of targets to be detected vary in size, and the use of a single-scale voxelization may result in excessive loss of detailed information, the voxels of different scales are introduced to extract relevant features of objects of different scales in the point clouds and integrate them into the second-stage detection. Lastly, a multi-scale feature fusion module is proposed, which simultaneously enhances and integrates features extracted from voxels of different scales. This module fully utilizes the valuable information present in the point cloud across various scales, ultimately leading to more precise 3D object detection. The experiment is conducted on the KITTI dataset and the nuScenes dataset. Compared with our baseline, "Pedestrian" detection improved by 3.37-2.72% and "Cyclist" detection by 3.79-1.32% across difficulty levels on KITTI, and was boosted by 2.4% in NDS and 3.6% in mAP on nuScenes.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001312992800001  
Indexed Date:  
2024-09-21  
Institution Address:  
North China Univ Technol, Sch Informat, Beijing 100144, Peoples R China  
Institution Address:  
Beijing Mass Transit Railway Operat Co Ltd, Corp Informat, Beijing 100044, Peoples R China  
Institution Address:  
Beijing Univ Chem Technol, Sch Informat, Beijing 100029, Peoples R China  
Institution:  
Institution:  
Beijing Mass Transit Railway Operat Co Ltd  
Institution:

### 727. DDFusion: An efficient multi-exposure fusion network with dense

摘要: In this work, we propose DDFusion, a novel multi-exposure image fusion network. DDFusion addresses the limitations of existing methods by effectively recovering details near extremely bright regions and learning associations between non-contiguous regions. To achieve this, our network incorporates a dense pyramidal (DensePy) convolution block in the encoder for multi-scale feature extraction, and a de-correlation fusion (DF) block for enabling structurally coherent and edge-preserving multi-scale feature fusion. It facilitates a smoother transition from highlighted areas to adjacent regions in the fused image. Experimental results demonstrate the superiority of DDFusion over state-of-the-art deep methods in terms of both visual quality and quantitative evaluation. Moreover, DDFusion achieves stronger multi-scale feature extraction capability with smaller computational complexity.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1047-3203  
eISSN:  
1095-9076  
入藏号:  
WOS:001099773000001  
Indexed Date:  
2023-11-26  
Institution Address:  
Shanghai Jiao Tong Univ, Sch Aeronaut & Astronaut, Shanghai 200240, Peoples R China  
Institution Address:  
Wuhan Second Ship Design & Res Inst, Wuhan 430010, Peoples R China  
Institution:  
Institution:  
Wuhan Second Ship Design & Res Inst

### 728. Scene Constrained Object Detection Method in High-Resolution

摘要: Ground objects in high-resolution remote sensing images are often closely related to the scenecategories. If the constraint information of the scene on the ground object can be usefully employed, it isexpected to improve further the performance of object detection. Considering the relationship between sceneinformation and objects, a scene constrained object detection method in high-resolution remote sensing imagesby Relation-aware Global Attention (RGA) is proposed. First, the global scene features are learned by addingthe global relational attention to the basic network in Feature fusion and Scaling-based Single Shot Detector(FS-SSD). Then, object is predicted by combining the oriented response convolution module with the multiscalefeature module under the constraints of learned global scene features. Finally, two loss functions are used tooptimize jointly the network to achieve object detection. Four experiments are conducted on NWPU VHR-10dataset and better object detection performance is achieved under the constraints of scene information  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
1009-5896  
入藏号:  
WOS:000860345900008  
Indexed Date:  
2022-10-15  
Institution Address:  
Beijing Univ Technol, Fac Informat Technol, Beijing 100124, Peoples R China  
Institution Address:  
Beijing Univ Technol, Beijing Key Lab Computat Intelligence & Intelligen, Beijing 100124, Peoples R China  
Institution Address:  
China Univ Geosci, Dept Resource Informat Engn, Wuhan 430074, Peoples R China  
Institution:  
Institution:  
Institution:

### 729. Cross-Connected Bidirectional Pyramid Network for Infrared Small-Dim

摘要: Infrared small-dim target detection is an important technology in the fields of infrared guidance, anti-missile, and tracking system. Due to the small size of targets, no obvious structure information, and low image signal-to-noise ratio (SNR), infrared small-dim target detection is still a challenging task. In this letter, a cross-connected bidirectional pyramid network (CBP-Net) is proposed for infrared small-dim target detection. The main body of the CBP-Net is to embed a bottom-up pyramid in the feature pyramid network (FPN), which is designed to provide more comprehensive target information by connecting with the original multi-scale features and the top-down pyramid. The bottom-up pyramid together with the top-down pyramid forms the proposed bidirectional pyramid structure. Then, an region of interest (ROI) feature augment module (RFA) composed of deformable ROI pooling and position attention is designed to fuse multi-scale ROI features and enhance the spatial information of the small-dim target. Besides, a regular constraint loss (RCL) is introduced to restrict multi-scale feature fusion to learn more precise target location information. Experimental results on two challenging datasets show that the performance of the proposed CBP-Net is superior to the state-of-the-art methods.  
Web of Science 核心合集中的 "被引频次":  
28  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
32  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000757847800002  
Indexed Date:  
2022-03-03  
Institution Address:  
Xidian Univ, Minist Educ, Key Lab Intelligent Percept & Image Understanding, Xian 710071, Peoples R China  
Institution Address:  
Xidian Univ, Acad Adv Interdisciplinary Res, Xian 710071, Peoples R China  
Institution Address:  
Dalian Maritime Univ, Sch Maritime Econ & Management, Dalian 116026, Peoples R China  
Institution Address:  
Chinese Acad Sci, Xian Inst Opt & Precis Mech, Xian 710119, Peoples R China  
Institution Address:  
Beijing Aerosp Automat Control Inst, Beijing 100070, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Beijing Aerosp Automat Control Inst

### 730. PanoDetNet: Multi-Resolution Panoramic Object Detection With Adaptive

摘要: Panoramic image object detection has significant applications in autonomous driving, robotic navigation, and security monitoring. However, most current object detection algorithms are trained on pinhole images and cannot be directly applied to panoramic images, which have a large field-of-view (FOV) and distortion. Additionally, research on panoramic image object detection lacks dedicated dataset support, and these images face challenges such as target distortion, occlusion, and multi-scale variations. Existing methods for panoramic image object detection have not yielded satisfactory performance. To address these issues, we propose PanoDetNet, an object detection model based on YOLOv7. We introduce two new modules: the Multi-Scale Feature Fusion (MSFF) module and the Adaptive Panoramic Feature Attention (APFA) module. The MSFF module enhances detection precision for targets of different scales by fusing feature maps of various sizes, while also reducing the number of parameters and simplifying the model structure. The APFA module adaptively addresses the distortion in panoramic images, improving the model's ability to locate and recognize objects in complex backgrounds and under occlusion. We trained PanoDetNet on our self-built panoramic image object detection dataset, PanoDet. This dataset was collected using a self-developed panoramic camera and manually annotated with the Labelme tool. Experimental results show that PanoDetNet achieves mAP@.5, mAP@.5:.95, and accuracy scores of 95.3%, 75.2%, and 94.1%, respectively. These results represent improvements of 1.6%, 2.5%, and 3.3% over YOLOv7. Our code is available at https://github.com/github98317/PanoDetNet.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001286667000001  
Indexed Date:  
2024-08-13  
Institution Address:  
Wuyi Univ, Fac Intelligent Mfg, Jiangmen 529000, Peoples R China  
Institution Address:  
Zhuhai 4Dage Network Technol, Zhuhai 519000, Peoples R China  
Institution:  
Institution:  
Zhuhai 4Dage Network Technol

### 731. Substation High-Voltage Switchgear Detection Based on Improved

摘要: Aiming at the current problems of low detection accuracy of high-voltage cabinet switches and large models that are difficult to deploy, a high-voltage cabinet switch detection method based on the improved EfficientNet-YOLOv5s model is proposed. Firstly, offline data expansion and online Mosaic data enhancement are performed based on the high-voltage switchgear images obtained by video frame extraction to alleviate the problems of unbalanced sample categories and the small number of samples. Secondly, the backbone network of the YOLOv5s model is replaced with a lightweight EfficientNet-B0 network, and the original convolution layer is replaced by depthwise separable convolution, and the original Squeeze-and-Excitation Networks were replaced with Efficient Channel Attention to optimize the efficiency and accuracy of the network model. To enhance the positioning information of small objects without bringing too much computational burden, the proposed improved multi-scale feature fusion Neck network uses the Bi-directional Concatenation(BiC) module to aggregate the feature maps of three adjacent layers. Finally, the H-Swish function is used to replace the Swish function in the MBConv structure, avoiding the exponential operation of the sigmoid function, and further improving the calculation efficiency of the model.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001216170100001  
Indexed Date:  
2024-05-13  
Institution Address:  
Zhejiang Sci Tech Univ, Sch Comp Sci & Technol, Hangzhou, Peoples R China  
Institution:

### 732. Real-Time Semantic Segmentation of Point Clouds Based on an Attention

摘要: A 3D point cloud is one of the main data sources for robot environmental cognition and understanding. Due to the limited computation and memory capacities of the robotic platform, existing semantic segmentation models of 3D point clouds cannot meet the requirements of real-time applications. To solve this problem, a lightweight, fully convolutional network based on an attention mechanism and a sparse tensor is proposed to better balance the accuracy and real-time performance of point cloud semantic segmentation. On the basis of the 3D-Unet structure, a global feature-learning module and a multi-scale feature fusion module are designed. The former improves the ability of features to describe important areas by learning the importance of spatial neighborhoods. The latter realizes the fusion of multi-scale semantic information and suppresses useless information through the task correlation learning of multi-scale features. Additionally, to efficiently process the large-scale point clouds acquired in real time, a sparse tensor-based implementation method is introduced. It is able to reduce unnecessary computation according to the sparsity of the 3D point cloud. As demonstrated by the results of experiments conducted with the SemanticKITTI and NuScenes datasets, our model improves the mIoU metric by 6.4% and 5%, respectively, over existing models that can be applied in real time. Our model is a lightweight model that can meet the requirements of real-time applications.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
eISSN:  
2076-3417  
入藏号:  
WOS:000947455100001  
Indexed Date:  
2023-03-30  
Institution Address:  
Dalian Maritime Univ, Coll Informat Sci & Technol, Dalian 116026, Peoples R China  
Institution:

### 733. MPA-GNet: multi-scale parallel adaptive graph network for 3D human pose

摘要: Graph convolutional networks (GCNs) have achieved remarkable performance in the 2D-to-3D human pose estimation (HPE) task. The adjacency matrix in GCNs is crucial for feature aggregation in 3D HPE. However, existing GCN-based methods excessively rely on the fixed adjacency matrix to aggregate joint features from one-hop neighbor at a single scale, which limits the feature representation of skeleton data. To better improve the performance of 3D HPE, we have designed a multi-scale parallel adaptive graph network (MPA-GNet) for 3D HPE. The proposed network consists of three parallel multi-scale subgraph networks (PMS-Net) to efficiently capture human joint features at different scales. Specially, a multi-scale feature fusion module is devised to process multi-scale graph structural features and exchange information to generate rich hierarchical representations for skeleton data. To flexible construct graph topology in different scales, a special designed adaptive attention adjacency graph convolution network and a cluster graph pooling module are designed to construct the MPA-GNet in a parallel manner and capture the local subgraphs information in each PMS-Net. Finally, we conduct experiments on two 3D human pose challenging benchmark datasets Human3.6M and HumanEva-I for evaluating the effectiveness of the proposed model. The experimental results demonstrate that our model achieves competitive performance compared with some state-of-the-art 3D HPE methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001099678500001  
Indexed Date:  
2023-11-28  
Institution Address:  
Shaanxi Normal Univ, Sch Comp Sci, Xian 710062, Peoples R China  
Institution Address:  
Shaanxi Normal Univ, Key Lab Modern Teaching Technol, Minist Educ, Xian 710062, Peoples R China  
Institution:  
Institution:

### 734. An Improved High Precision 3D Semantic Mapping of Indoor Scenes from

摘要: This paper proposes an improved high-precision 3D semantic mapping method for indoor scenes using RGB-D images. The current semantic mapping algorithms suffer from low semantic annotation accuracy and insufficient real-time performance. To address these issues, we first adopt the Elastic Fusion algorithm to select key frames from indoor environment image sequences captured by the Kinect sensor and construct the indoor environment space model. Then, an indoor RGB-D image semantic segmentation network is proposed, which uses multi-scale feature fusion to quickly and accurately obtain object labeling information at the pixel level of the spatial point cloud model. Finally, Bayesian updating is used to conduct incremental semantic label fusion on the established spatial point cloud model. We also employ dense conditional random fields (CRF) to optimize the 3D semantic map model, resulting in a high-precision spatial semantic map of indoor scenes. Experimental results show that the proposed semantic mapping system can process image sequences collected by RGB-D sensors in real-time and output accurate semantic segmentation results of indoor scene images and the current local spatial semantic map. Finally, it constructs a globally consistent high-precision indoor scenes 3D semantic map.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1526-1492  
eISSN:  
1526-1506  
入藏号:  
WOS:001021538000001  
Indexed Date:  
2023-07-16  
Institution Address:  
Xian Univ Technol, Shaanxi Key Lab Complex Syst Control & Intelligent, Xian 710048, Peoples R China  
Institution Address:  
Univ Sydney, Australian Ctr Field Robot, Sydney 2006, Australia  
Institution:  
Institution:

### 735. Real-time automated detection of multi-category solar radio bursts

摘要: Accurate real-time solar radio burst (SRB) detection is crucial for solar physics research and space weather forecasting. Currently, most studies on solar radio burst detection focus on single-category identification and simple discrimination of bursts. There are limited existing studies on multi-category detection. This paper proposes a real-time multi-category solar radio burst detection method to meet the requirements of real-time detection, detection accuracy, and classification accuracy in solar radio bursts. First, solar radio burst spectrums were collected from e-CALLISTO. The spectrums are labeled using LabelImg, and a dataset containing solar radio bursts of Type II, Type III, Type IIIs, Type IV, and Type V was established. Second, a full-dimensional dynamic convolution was introduced in the backbone module of the YOLOv8n model, enhancing the model's feature extraction capability. Third, a multi-scale feature fusion network based on ConvNeXt was created to prevent feature information loss and optimize the loss function. The experimental results show that the proposed method achieves an average detection accuracy of 82.4% on the established solar radio burst dataset. Compared with the original YOLOv8n model, the accuracy increased by 3.5%. Additionally, the model operates at 140.9 frames per second, with each frame representing a spectrum of 15 minutes duration. Thus, the improved YOLOv8n model enhances the detection accuracy and speed of solar radio bursts, enabling automatic detection and localization of solar radio bursts of Type II, Type III, Type IIIs, Type IV, and Type V.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
0004-640X  
eISSN:  
1572-946X  
入藏号:  
WOS:001324565000001  
Indexed Date:  
2024-10-07  
Institution Address:  
Yunnan Univ, Sch Informat Sci & Engn, Kunming 650504, Peoples R China  
Institution Address:  
Chinese Acad Sci, State Key Lab Space Weather, Beijing 100190, Peoples R China  
Institution Address:  
Yunnan Key Lab Intelligent Syst & Comp, Kunming 650504, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Astron & Space Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:  
Institution:  
Yunnan Key Lab Intelligent Syst & Comp  
Institution:

### 736. Research on Deep Learning Model Enhancements for PCB Surface Defect

摘要: With the miniaturization and increasing complexity of electronic devices, the accuracy and efficiency of printed circuit board (PCB) defect detection are crucial to ensuring product quality. To address the issues of small defect sizes and high missed detection rates in PCB surface inspection, this paper proposes an enhanced YOLOv8s model which not only improves detection performance but also achieves a lightweight design. Firstly, the Nexus Attention module is introduced, which organically integrates multiple attention mechanisms to further enhance feature extraction and fusion capabilities, improving the model's learning and generalization performance. Secondly, an improved CGFPN network is designed to optimize multi-scale feature fusion, significantly boosting the detection of small objects. Additionally, the WaveletUnPool module is incorporated, leveraging wavelet transform technology to refine the upsampling process, accurately restoring detailed information and improving small-object detection in complex backgrounds. Lastly, the C2f-GDConv module replaces the traditional C2f module, reducing the number of model parameters and computational complexity while maintaining feature extraction efficiency. Comparative experiments on a public PCB dataset demonstrate that the enhanced model achieved a mean average precision (mAP) of 97.3% in PCB defect detection tasks, representing a 3.0% improvement over the original model, while reducing Giga Floating Point Operations (GFLOPs) by 26.8%. These enhancements make the model more practical and adaptable for industrial applications, providing a solid foundation for future research.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001377927200001  
Indexed Date:  
2024-12-25  
Institution Address:  
Fujian Polytech Normal Univ, Key Lab Nondestruct Testing, Fuqing 350300, Peoples R China  
Institution Address:  
Fujian Polytech Normal Univ, Sch Elect & Mech Engn, Fuqing 350300, Peoples R China  
Institution:  
Institution:

### 737. Retouched Face Image Quality Assessment Based on Differential Perception

摘要: Face retouching involves using digital techniques to alter an individual's appearance, commonly using in social media. However, excessively retouched face (RF) images can lead to issues such as unrealistic beauty standards and psychological stress. Therefore, it is crucial to develop a reliable quality assessment method for RF images. In this paper, we propose a novel network named DIRF-IQA for RF image quality assessment (IQA). DIRF-IQA mainly includes a parameter-shared image encoder, a text encoder, and three key components, namely the Differential Feature Attention Module (DFAM), the Text-image Interaction Module (TIM), and the Multi-scale Feature Fusion Module (MFFM). Specifically, the DFAM captures both local and global differences between original and retouched images by processing multi-scale features and utilizing cross-attention and self-attention blocks for differential perception. In the TIM, textual prompts summarizing retouching operations are encoded by a text encoder and integrated with differential features extracted by the DFAM to enhance the understanding of distortions in RF images. The MFFM then fuses these text-enhanced features across different layers and combines them with the global differential feature to predict the quality of the retouched images. We conduct extensive experiments on two RF IQA databases and the results demonstrate the superiority of DIDF-IQA compared to 12 state-of-the-art full-reference IQA methods in evaluating RF images.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0018-9316  
eISSN:  
1557-9611  
入藏号:  
WOS:001308200700001  
Indexed Date:  
2024-09-18  
Institution Address:  
Shenzhen Univ, Sch Management, Shenzhen 518060, Peoples R China  
Institution Address:  
Baicheng Normal Univ, Sch Mech & Control Engn, Baicheng 137000, Peoples R China  
Institution Address:  
Shenzhen Univ, Coll Elect & Informat Engn, Guangdong Key Lab Intelligent Informat Proc, Shenzhen 518060, Peoples R China  
Institution Address:  
China Univ Geosci, Sch Comp Sci, Wuhan 430074, Peoples R China  
Institution Address:  
Shenzhen MSU BIT Univ, Guangdong Lab Machine Percept & Intelligent Comp, Shenzhen 518115, Peoples R China  
Institution Address:  
Shenzhen Univ, Sch Biomed Engn, Shenzhen 518060, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 738. An Attention-Based Full-Scale Fusion Network for Segmenting Roof Mask

摘要: Accurately segmenting building roofs from satellite images is crucial for evaluating the photovoltaic power generation potential of urban roofs and is a worthwhile research topic. In this study, we propose an attention-based full-scale fusion (AFSF) network to segment a roof mask from the given satellite images. By developing an attention-based residual ublock, the channel relationship of the feature maps can be modeled. By integrating attention mechanisms in multi-scale feature fusion, the model can learn different weights for features of different scales. We also design a ladder-like network to utilize weakly labeled data, thereby achieving pixel-level semantic segmentation tasks assisted by image-level classification tasks. In addition, we contribute a new roof segmentation dataset, which is based on satellite images and uses the roof as the segmentation target rather than the entire building to further promote the algorithm research of estimating roof area using satellite images. The experimental results on the new roof segmentation dataset, WHU dataset, and IAIL dataset demonstrate the effectiveness of the proposed network.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001245776800001  
Indexed Date:  
2024-06-21  
Institution Address:  
Tsinghua Univ, Inst Future Human Habitats, Tsinghua Shenzhen Int Grad Sch, Shenzhen 518055, Peoples R China  
Institution Address:  
Harbin Inst Technol Shenzhen, Shenzhen Grad Sch, Shenzhen 518055, Peoples R China  
Institution Address:  
Natl Supercomuting Ctr Shenzhen, Shenzhen 518055, Peoples R China  
Institution Address:  
CCSTC China Construct Sci & Technol Cooperat Low C, Beijing 100195, Peoples R China  
Institution:  
Institution:  
Institution:  
Natl Supercomuting Ctr Shenzhen  
Institution:  
CCSTC China Construct Sci & Technol Cooperat Low C

### 739. DM-YOLOX aerial object detection method with intensive attention

摘要: In aerial image detection, difficulties in feature extraction and low detection accuracy arise due to background interference, occlusion, and the presence of multiple small objects. This paper proposes a DM-YOLOX aerial object target detection method with intensive attention mechanism. Firstly, the proposed approach incorporates coordinate attention (CA) and a dense connection method into the backbone network architecture, enabling adaptive channel weighting throughout the feature extraction process. This facilitates the enhancement of significant features while suppressing less relevant ones, thereby augmenting the network's capacity to represent object features and ensuring retention and reinforcement of key features. Secondly, the multibranch extraction module (MBE) is incorporated into the feature fusion network to enhance the network's ability in extracting multi-scale feature information from images with extensive coverage, thereby enhancing the detection accuracy and efficiency of small- and medium-sized objects in complex scenes. Finally, the utilization of SIoU instead of IoU as the bounding box loss function effectively addresses the issue of mismatch between real and predicted boxes, leading to accelerated network convergence and improved performance during model training. After training and testing on the VisDrone 2019 dataset, this method effectively detects small objects in complex environments. The DM-YOLOX model shows a significant improvement of 2.7% in mAP compared to the baseline network, while achieving an 8% increase in frames per second (FPS).  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
0920-8542  
eISSN:  
1573-0484  
入藏号:  
WOS:001166424800001  
Indexed Date:  
2024-03-04  
Institution Address:  
Xian Polytech Univ, Xian 710600, Shaanxi, Peoples R China  
Institution:

### 740. FQTrack:Object Tracking Method Based on a Feature-Enhanced Memory

摘要: Visual object tracking technology is widely used in intelligent security, automatic driving and other fields, and also plays an important role in frontier fields such as human-computer interactions and virtual reality. The memory network improves the stability and accuracy of tracking by using historical frame information to assist in the positioning of the current frame in object tracking. However, the memory network is still insufficient in feature mining and the accuracy and robustness of the model may be reduced when using noisy observation samples to update it. In view of the above problems, we propose a new tracking framework, which uses the attention mechanism to establish a feature-enhanced memory network and combines cross-attention to aggregate the spatial and temporal context information of the target. The former introduces spatio-temporal adaptive attention and cross-spatial attention, embeds spatial location information into channels, realizes multi-scale feature fusion, dynamically emphasizes target location information, and obtains richer feature maps. The latter guides the tracker to focus on the area with the largest amount of information in the current frame to better distinguish the foreground and background. In addition, through the memory quality selection mechanism, the accuracy and richness of the feature samples are improved, thereby enhancing the adaptability and discrimination ability of the tracking model. Experiments on benchmark test sets such as OTB2015, TrackingNet, GOT-10k, LaSOT and UAV 123 show that this method achieves comparable performance with advanced trackers.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001305864600001  
Indexed Date:  
2024-09-15  
Institution Address:  
Zhengzhou Univ Light Ind, Sch Software, Zhengzhou 450000, Peoples R China  
Institution Address:  
Zhengzhou Univ Light Ind, Sch Elect & Informat Engn, Zhengzhou 450000, Peoples R China  
Institution Address:  
Zhengzhou Univ Light Ind, Sch Comp & Commun Engn, Zhengzhou 450000, Peoples R China  
Institution:  
Institution:  
Institution:

### 741. BDK-YOLOv8: An Enhanced Algorithm for UAV Infrared Image Object

摘要: This paper presents an infrared small object detection algorithm based on YOLOv8n to address challenges like large model size, complex backgrounds, poor small object detection, and scale variations. First, a new C2f-DCNv3 module is introduced to reduce parameter redundancy and enhance feature extraction. A Bidirectional Feature Pyramid Network (BiFPN) is added to the neck structure for improved detection of very small objects, enabling better multi-scale feature fusion. An improved SIOU loss function is also proposed, prioritizing small object samples and those with average-quality annotations. Finally, channel pruning is applied to reduce model parameters and computational complexity, improving detection efficiency.Experimental results show that the proposed algorithm achieves 94.3% mAP50 on the HIT-UAV dataset, a 1.6% improvement over the original YOLOv8n, with a 3% increase in recall. Model parameters and computational load are reduced by 55.1% and 1.2%, respectively, while the model size decreases by 1.77MB. Overall, the improved model offers a strong balance between accuracy and efficiency, making it well-suited for embedded devices and industrial drone detection in various scenarios.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001381662100010  
Indexed Date:  
2024-12-27  
Institution Address:  
Nanchang Univ, Sch Informat Engn, Nanchang 330031, Peoples R China  
Institution:

### 742. Detection algorithm of aircraft skin defects based on improved YOLOv8n

摘要: In order to solve the problem of small targets being prone to false detection and missed detection in aircraft skin defect detection under complex backgrounds, the model of aircraft skin defect detection based on improved YOLOv8n is proposed in this paper. Firstly, the Shuffle Attention + + module is incorporated into the network, combined with the residual connection idea, to more efficiently fuse feature map information; Secondly, SIOU and Focal Loss are used to replace CIOU as the regression loss functions to balance positive and negative samples in complex backgrounds and accelerate model convergence; Subsequently, the bidirectional feature pyramid network is used to modify the detection head and enhance multi-scale feature fusion. Furthermore, the depth-wise convolution module is used to replace the convolution module (Conv) in the neck part, which serves to reduce the parameters of the model and speed up the detection speed. Finally, an aircraft skin defect dataset is established, combined with Mosaic data enhancement to prevent the model from overfitting, and adopted the class balancing strategy to avoid class bias. The experimental results show that the detection accuracy of our improved YOLOv8n model is 97.9%, which is 7.3% higher than the baseline model. The model's recall rate, the mean average precision, and F1 scores are improved by 13.9%, 6.6%, and 11.0%, respectively. The detection speed has achieved 139FPS, fulfilling the requirements of high accuracy and real-time performance in small target aircraft skin defect detection tasks.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001171884200003  
Indexed Date:  
2024-03-16  
Institution Address:  
Civil Aviat Univ China, Engn Tech Training Ctr, Tianjin 300300, Peoples R China  
Institution Address:  
Civil Aviat Univ China, Coll Aeronaut Engn, Tianjin 300300, Peoples R China  
Institution Address:  
Civil Aviat Univ China, Inst Sci & Technol Innovat, Tianjin 300300, Peoples R China  
Institution:  
Institution:  
Institution:

### 743. A Cross-Level Iterative Subtraction Network for Camouflaged Object

摘要: Camouflaged object detection (COD) is a challenging task, aimed at segmenting objects that are similar in color and texture to their background. Sufficient multi-scale feature fusion is crucial for accurately segmenting object regions. However, most methods usually focus on information compensation, overlooking the difference between features, which is important for distinguishing the object from the background. To this end, we propose the cross-level iterative subtraction network (CISNet), which integrates information from cross-layer features and enhances details through iteration mechanisms. CISNet involves a cross-level iterative structure (CIS) for feature complementarity, where texture information is used to enrich high-level features and semantic information is used to enhance low-level features. In particular, we present a multi-scale strip convolution subtraction (MSCSub) module within CIS to extract difference information between cross-level features and fuse multi-scale features, which improves the feature representation and guides accurate segmentation. Furthermore, an enhanced guided attention (EGA) module is presented to refine features by deeply mining local context information and capturing a broader range of relationships between different feature maps in a top-down manner. Extensive experiments conducted on four benchmark datasets demonstrate that our model outperforms the state-of-the-art COD models in all evaluation metrics.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001311355400001  
Indexed Date:  
2024-09-19  
Institution Address:  
Hohai Univ, Coll Comp Sci & Software Engn, Nanjing 211100, Peoples R China  
Institution Address:  
Minist Water Resources, Informat Ctr, Beijing 100053, Peoples R China  
Institution Address:  
Hohai Univ, Key Lab Water Big Data Technol, Minist Water Resources, Nanjing 211100, Peoples R China  
Institution Address:  
Water Resources Serv Ctr Jiangsu Prov, Nanjing 210029, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Water Resources Serv Ctr Jiangsu Prov

### 744. Hyperspectral image classification Based on weakened Laplacian pyramid

摘要: In recent years, methods of spatial-spectral feature extraction have been widely used in hyperspectral image (HSI) classification and have achieved good performance. However, the ways are often unable to distinguish effectively the boundary of the adjacent ground objects. Multi-scale feature fusion alleviates the problems to a certain extent, but in the case of limited samples, the previous methods often fail to achieve the desired effect. In this paper, we propose a multi-scale hyperspectral image classification method based on weakened Laplacian pyramid and guided filtering to nicely distinguish the boundary of the different ground objects. This structure first obtains multi-scale information by constructing a weakened Laplacian pyramid, and weakens the magnification of Gaussian blur in the upsampling process. While obtaining multi-scale images, it can not only retain enough image information to expand feature samples but also emphasize edge information in the edge extraction of the Laplacian pyramid. Then, the weakened Laplacian pyramid feature map is used as the guiding image of the guiding filter, which can better preserve the edge features and further improve the classification performance of HSI. Extensive experiments are carried out on four HSI data sets. The results show that our method is superior in classification accuracy and computational cost.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:001058070600001  
Indexed Date:  
2023-09-14  
Institution Address:  
China Univ Petr East China, Coll Oceanog & Space Informat, Qingdao, Peoples R China  
Institution Address:  
China Univ Petr East China, Coll Control Sci & Engn, Qingdao, Peoples R China  
Institution Address:  
Minist Nat Resources, Inst Oceanog 1, Lab Marine Phys & Remote Sensing, Qingdao, Peoples R China  
Institution:  
Institution:  
Institution:

### 745. IPDC Network-Based Image Reconstruction for Damage Detection of CFRP

摘要: The nonlinearity and ill-poseness of image reconstruction has hindered the application of electrical impedance tomography (EIT) in damage detection of Carbon Fiber-Reinforced Polymer (CFRP). To cope with this problem, a dual-channel Inception-Dense-Cbam (IPDC) network model integrated with attention mechanism is proposed. To combine the diversity of feature extraction with the sparseness of the true conductivity distribution, the whole frame is designed to be composed of two parallel branches: the Inception-resnet-cbam (IPRC) branch utilizing multi-scale feature fusion based on the Inception-resnet module and the Dense-Cbam (Dense-CB) branch utilizing feature reuse based on the dense module. Features extracted via the two branches are then fused as outputs. Additionally, to ensure extraction accuracy of damage boundary, convolutional block attention mechanism module (CBAM) is integrated into both branches. Both simulation results and prototype experiments showed that compared with traditional algorithms, the proposed method could effectively reduce the artifacts of damage images, improve accuracy of damage location, and enhance the definition of damage edges. Simulation data obtained with added noise and the data of prototype experiments demonstrated that the IPDC network had the good anti-noise performance in detecting different types of damages.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
8  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:000819823500089  
Indexed Date:  
2022-07-14  
Institution Address:  
Civil Aviat Univ China, Sch Elect Informat & Automat, Tianjin 300300, Peoples R China  
Institution:

### 746. Application Research of Bridge Damage Detection Based on the Improved

摘要: To ensure the safety and rational use of bridge traffic lines, the existing bridge structural damage detection models are not perfect for feature extraction and have difficulty meeting the practicability of detection equipment. Based on the YOLO (You Only Look Once) algorithm, this paper proposes a lightweight target detection algorithm with enhanced feature extraction of bridge structural damage. The BIFPN (Bidirectional Feature Pyramid Network) network structure is used for multi-scale feature fusion, which enhances the ability to extract damage features of bridge structures, and uses EFL (Equalized Focal Loss) to optimize the sample imbalance processing mechanism, which improves the accuracy of bridge structure damage target detection. The evaluation test of the model has been carried out in the constructed BDD (Bridge Damage Dataset) dataset. Compared with the YOLOv3-tiny, YOLOv5S, and B-YOLOv5S models, the mAP@.5 of the BE-YOLOv5S model increased by 45.1%, 2%, and 1.6% respectively. The analysis and comparison of the experimental results prove that the BE-YOLOv5S network model proposed in this paper has a better performance and a more reliable performance in the detection of bridge structural damage. It can meet the needs of bridge structure damage detection engineering with high requirements for real-time and flexibility.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
eISSN:  
2076-3417  
入藏号:  
WOS:000815983300001  
Indexed Date:  
2022-07-01  
Institution Address:  
Ocean Univ China, Coll Engn, Dept Civil Engn, Qingdao 266100, Peoples R China  
Institution:

### 747. Learning Unoccluded Face Texture Completion from Single Image in the

摘要: In recent years, single tasks such as face frontalization, image inpainting, and glasses removal have improved face de-occlusion. However, there is little work on joint learning of multiple de-occlusion tasks. To achieve multi-task learning, we propose an unoccluded face synthesis (UFS) framework for multi-tasks such as face frontalization, image inpainting, and glasses removal, which can remove glasses, face self-occlusion, and external occlude. Our UFS framework consists of an encoder, an image reconstruction module, a decoder, and an image discriminator. First, Gaussian random noise extracts high-dimensional features from images in the encoder module. Next, the image reconstruction module includes multi-scale feature fusion, residual hole block, and self-attention network. As a result, it can strengthen the learning of multi-level fine-grained features and achieve better results in face restoration and face frontalization tasks. Then, we synthesize unoccluded face textures from multi-level fine-grained elements in the decoder. Finally, the image discriminator learns the global information structure of the synthesized image, preventing problems such as distortion and blurring of the picture. Experiments show that our UFS framework can achieve better results on single tasks such as face frontalization, image inpainting, and glasses removal. It also can obtain acceptable results on multiple tasks such as face frontalization and glasses removal simultaneously.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1370-4621  
eISSN:  
1573-773X  
入藏号:  
WOS:000795757600001  
Indexed Date:  
2022-05-25  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao 266590, Peoples R China  
Institution Address:  
Beihang Univ, State Key Lab Virtual Real Technol & Syst, Beijing 100191, Peoples R China  
Institution Address:  
Beihang Univ, Virtual Real Res Inst, Qingdao Res Inst, Qingdao 266100, Peoples R China  
Institution:  
Institution:  
Institution:

### 748. Multi-Scale Feature Aggregation Network for Water Area Segmentation

摘要: Water area segmentation is an important branch of remote sensing image segmentation, but in reality, most water area images have complex and diverse backgrounds. Traditional detection methods cannot accurately identify small tributaries due to incomplete mining and insufficient utilization of semantic information, and the edge information of segmentation is rough. To solve the above problems, we propose a multi-scale feature aggregation network. In order to improve the ability of the network to process boundary information, we design a deep feature extraction module using a multi-scale pyramid to extract features, combined with the designed attention mechanism and strip convolution, extraction of multi-scale deep semantic information and enhancement of spatial and location information. Then, the multi-branch aggregation module is used to interact with different scale features to enhance the positioning information of the pixels. Finally, the two high-performance branches designed in the Feature Fusion Upsample module are used to deeply extract the semantic information of the image, and the deep information is fused with the shallow information generated by the multi-branch module to improve the ability of the network. Global and local features are used to determine the location distribution of each image category. The experimental results show that the accuracy of the segmentation method in this paper is better than that in the previous detection methods, and has important practical significance for the actual water area segmentation.  
Web of Science 核心合集中的 "被引频次":  
44  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
44  
eISSN:  
2072-4292  
入藏号:  
WOS:000758649700001  
Indexed Date:  
2022-03-04  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Collaborat Innovat Ctr Atmospher Environm & Equip, B DAT, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Forestry Univ, Coll Informat Sci & Technol, Nanjing 210037, Peoples R China  
Institution:  
Institution:

### 749. Multi-scale differential network for landslide extraction from remote

摘要: Landslides are major geological hazards globally, causing significant economic losses each year. Accurate landslide detection is essential for disaster prevention, risk assessment, and timely emergency response. Current extraction methods struggle to distinguish landslides from their surroundings and precisely define their boundaries. To address these challenges, we introduce the Multi-Scale Difference Enhancement Network (MSDENet), a framework for landslide extraction through time-based change detection. MSDENet incorporates three core components: the Difference Guided Attention Module (DGAM) for enhanced focus on landslide-specific changes, the Multi-Scale Feature Fusion Module (MSFFM) for improved boundary delineation, and the Multi-Scale Sensory Module (MSSM) to boost generalization by integrating multi-scale features. We validate MSDENet's effectiveness on the Global Very-High-Resolution Landslide Mapping (GVLM) dataset, covering 17 diverse landslide events, and further assess its applicability on high-resolution Nepal and Wenchuan datasets. MSDENet outperforms six contemporary frameworks, achieving IoU improvements of 1.42% and 1.08% for the Kaikoura and Tbilisi datasets and demonstrating gains of 3.97% and 4.79% for the Nepal and Wenchuan datasets, confirming its effectiveness in varied conditions.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1753-8947  
eISSN:  
1753-8955  
入藏号:  
WOS:001379797200001  
Indexed Date:  
2024-12-24  
Institution Address:  
Int Res Ctr Big Data Sustainable Dev Goals, Beijing 100094, Peoples R China  
Institution Address:  
Guilin Univ Elect Technol, Sch Comp Sci & Informat Secur, Guilin, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Key Lab Digital Earth Sci, Beijing 100094, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Atmospher Phys, Beijing, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 750. Video Object Segmentation Using Multi-Scale Attention-Based Siamese

摘要: Video target segmentation is a fundamental problem in computer vision that aims to segment targets from a background by learning their appearance information and movement information. In this study, a video target segmentation network based on the Siamese structure was proposed. This network has two inputs: the current video frame, used as the main input, and the adjacent frame, used as the auxiliary input. The processing modules for the inputs use the same structure, optimization strategy, and encoder weights. The input is encoded to obtain features with different resolutions, from which good target appearance features can be obtained. After processing using the encoding layer, the motion features of the target are learned using a multi-scale feature fusion decoder based on an attention mechanism. The final predicted segmentation results were calculated from a layer of decoded features. The video object segmentation framework proposed in this study achieved optimal results on CDNet2014 and FBMS-3D, with scores of 78.36 and 86.71, respectively. It outperformed the second-ranked method by 4.3 on the CDNet2014 dataset and by 0.77 on the FBMS-3D dataset. Suboptimal results were achieved on the video primary target segmentation datasets SegTrackV2 and DAVIS2016, with scores of 60.57 and 81.08, respectively.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2079-9292  
入藏号:  
WOS:001028636700001  
Indexed Date:  
2023-07-28  
Institution Address:  
East China Jiaotong Univ, Sch Software, Nanchang 330013, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Software, State Key Lab Comp Sci, Beijing 100190, Peoples R China  
Institution Address:  
Nanchang Normal Univ, Dept Math & Comp Sci, Nanchang 330032, Peoples R China  
Institution:  
Institution:  
Institution:

### 751. Improvement of Lightweight Convolutional Neural Network Model Based on

摘要: To ensure the safe operation of highway traffic lines, given the imperfect feature extraction of existing road pit defect detection models and the practicability of detection equipment, this paper proposes a lightweight target detection algorithm with enhanced feature extraction based on the YOLO (You Only Look Once) algorithm. The BIFPN (Bidirectional Feature Pyramid Network) network structure is used for multi-scale feature fusion to enhance the feature extraction ability, and Varifocal Loss is used to optimize the sample imbalance problem, which improves the accuracy of road defect target detection. In the evaluation test of the model in the constructed PCD1 (Pavement Check Dataset) dataset, the mAP@.5 (mean Average Precision when IoU = 0.5) of the BV-YOLOv5S (BiFPN Varifocal Loss-YOLOv5S) model increased by 4.1%, 3%, and 0.9%, respectively, compared with the YOLOv3-tiny, YOLOv5S, and B-YOLOv5S (BiFPN-YOLOv5S; BV-YOLOv5S does not use the Improved Focal Loss function) models. Through the analysis and comparison of experimental results, it is proved that the proposed BV-YOLOv5S network model performs better and is more reliable in the detection of pavement defects and can meet the needs of road safety detection projects with high real-time and flexibility requirements.  
Web of Science 核心合集中的 "被引频次":  
55  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
59  
eISSN:  
1424-8220  
入藏号:  
WOS:000796091600001  
PubMed ID:  
35591227  
Indexed Date:  
2022-05-25  
Institution Address:  
Ocean Univ China, Dept Civil Engn, Coll Engn, Qingdao 266100, Peoples R China  
Institution:

### 752. Adaptive Circular Receptive Field for Remote Sensing Image Target

摘要: Remote sensing images contain a variety of targets. The circular and square-like targets are very common. However, general methods are prone to incorporating background information and have problems with rotation angles when detecting circular and square-like targets. The Adaptive Circular Receptive Field Network (ACRFNet) is proposed to address these problems. Firstly, a radius detection head based on the center point is designed, which predicts the center point and radius of the target to obtain a bounding circle. Compared with the rectangular bounding box, the bounding circle contains less background information and has rotation invariance, achieving in more accurate localization. Secondly, in order to better extract the features of circular targets and cope with scale changes, a multi-scale feature fusion method based on adaptive circular convolution is proposed. Adaptive circular receptive fields replace the ordinary convolution of square receptive fields, and a feature pyramid structure is introduced to enrich the semantic information of the feature map. Finally, in order to reasonably evaluate the performance of the model, Circle IOU is used to calculate mAP to fit the bounding circle. The experimental results on the NWPU VHR-10 and DIOR datasets demonstrate that the model has significantly improved detection accuracy for circular and square-like targets and has better generalization for other types of targets.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001425531400003  
Indexed Date:  
2025-02-26  
Institution Address:  
Nanchang Hangkong Univ, Sch Software, Nanchang 330063, Jiangxi, Peoples R China  
Institution Address:  
Nanchang Hangkong Univ, Sch Civil Engn & Architecture, Nanchang 330063, Jiangxi, Peoples R China  
Institution Address:  
Huihang Jiangxi Digital Technol Co Ltd, Nanchang 330038, Jiangxi, Peoples R China  
Institution:  
Institution:  
Institution:  
Huihang Jiangxi Digital Technol Co Ltd

### 753. Low Light Image Enhancement Based on Multi-Scale Network Fusion

摘要: At present, researchers have made great progress in the research of object detection, however, these studies mainly focus on the object detection of images under normal lighting, ignoring the target detection under low light. And images in the fields of automatic driving at night and surveillance are usually obtained in low-light environments. These images have problems such as poor brightness, low contrast, and obvious noise, which lead to a large amount of information loss in the image. And the performance of object detection in low light is reduced. In this paper, we propose a low-light image enhancement method based on multi-scale network fusion to solve the problems of images in low-light environments. Aiming at the problem that the effective information of low-light images is relatively small, we propose a preprocessing method for image nonlinear transformation and fusion, which improves the amount of available information in the light image. Then, in order to obtain a better enhancement effect, a multi-scale feature fusion method is proposed, which fuses features from different resolution levels in the network. The details of low-light areas in the image are improved, and the problem of feature loss caused by too deep network layers is solved. The experimental results show that our proposed method can achieve better enhancement effects on different datasets compared with the current mainstream methods. The average recall value of the object detection with our method is improved by 38.25%, which shows that our proposed method is effective and can promote the development of autonomous driving, monitoring, and other fields.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
2169-3536  
入藏号:  
WOS:000899137300001  
Indexed Date:  
2023-01-01  
Institution Address:  
Changchun Univ Sci & Technol, Sch Elect & Informat Engn, Changchun 130022, Peoples R China  
Institution Address:  
Changchun Univ Sci & Technol, Sch Optoelect Engn, Changchun 130022, Peoples R China  
Institution:  
Institution:

### 754. Citrus Disease Classification Model Based on Improved ConvNeXt

摘要: Early diagnosis of citrus diseases directly affects the yield and quality of citrus cultivation, and a citrus disease classification model based on improved ConvNeXt is proposed to address the problems of high cost and low efficiency of traditional citrus disease detection methods. Firstly, the attention mechanism is changed to a parallel connection in the convolutional block attention module, and the attention mechanism is incorporated into ConvNeXt, which improves the model's ability of feature extraction, makes the model focus more on lesion features, and suppresses the interference of background information. Secondly, the multi-scale feature fusion module is incorporated to improve the model's adaptability to disease features at different scales and improve the network classification performance. Finally, the transfer learning method is used to conduct pre-training with ImageNet weight information to reduce the impact of insufficient samples. The experimental results show that on the self-constructed complex background citrus disease dataset, the model in this paper achieves an average accuracy of 98.07%, which is improved by 2.9% compared with the original model, and the comprehensive performance is significantly better than that of VGG, AlexNet, and other models. The model in this paper improves the performance of disease classification and provides a theoretical basis for intelligent classification of citrus diseases.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001340696500001  
Indexed Date:  
2024-10-30  
Institution Address:  
Guilin Univ Technol, Key Lab Adv Mfg & Automat Technol, Educ Dept Guangxi Zhuang Autonomous Reg, Guilin 541006, Peoples R China  
Institution Address:  
Guilin Univ Technol, Sch Mech & Control Engn, Guilin 541006, Peoples R China  
Institution Address:  
Guilin Univ Technol, Sch Surveying Mapping & Geog Informat, Guilin 541006, Peoples R China  
Institution Address:  
Jimei Univ, Sch Marine Engn, Xiamen 361021, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 755. YOLO-SAD: An Efficient SAR Aircraft Detection Network

摘要: Aircraft detection in SAR images of airports remains crucial for continuous ground observation and aviation transportation scheduling in all weather conditions, but low resolution and complex scenes pose unique challenges. Existing methods struggle with accuracy, overlapping detections, and missed targets. We propose You Only Look Once-SAR Aircraft Detector (YOLO-SAD), a novel detector that tackles these issues. YOLO-SAD leverages the Attention-Efficient Layer Aggregation Network-Head (A-ELAN-H) module to prioritize essential features for improved accuracy. Additionally, the SAR Aircraft Detection-Feature Pyramid Network (SAD-FPN) optimizes multi-scale feature fusion, boosting detection speed. Finally, Enhanced Non-Maximum Suppression (EH-NMS) eliminates overlapping detections. On the SAR Aircraft Detection Dataset (SADD), YOLO-SAD achieved 91.9% AP(0.5) and 57.1% AP(0.5:0.95), surpassing the baseline by 2.1% and 1.9%, respectively. Extensive comparisons on SADD further demonstrate YOLO-SAD's superiority over five state-of-the-art methods in both AP(0.5) and AP(0.5:0.95). The outcomes of further comparative experiments on the SAR-AIRcraft-1.0 dataset confirm the robust generalization capability of YOLO-SAD, demonstrating its potential use in aircraft detection with SAR.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
eISSN:  
2076-3417  
入藏号:  
WOS:001201199200001  
Indexed Date:  
2024-04-17  
Institution Address:  
Sun Yat Sen Univ, Sch Aeronaut & Astronaut, Shenzhen Campus,66 Gongchang Rd, Shenzhen 518107, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Shenzhen Key Lab Intelligent Microsatellite Conste, Shenzhen Campus,66 Gongchang Rd, Shenzhen 518107, Peoples R China  
Institution:  
Institution:

### 756. Real-time and effective pan-sharpening for remote sensing using

摘要: Real-time monitoring and surveillance play an important role in the field of remote sensing, where multi-spectral (MS) images with high spatial resolution are widely desired for better analysis. However, high-resolution MS images cannot be directly obtained due to the limitations of sensors and bandwidth. As an essential way to alleviate this problem, pan-sharpening aims at fusing the complementary information of a low-resolution MS image and a high-resolution panchromatic (PAN) image to reconstruct a high-resolution MS image. Most previous deep-learning based methods can meet the real-time requirements with the help of graphics processing unit (GPU). However, they don't fully exploit the favorable hierarchical information, sparing huge room for performance improvement. In this paper, to meet the requirement of real-time implementation and achieve more effective performance simultaneously, we propose a multi-scale fusion network (MSFN) to make full use of hierarchical complementary features of PAN and MS images. Specifically, we introduce an encoder-decoder structure and coarse-to-fine strategy to effectively extract multi-scale features of PAN and MS images, separately. Meanwhile, an information pool is adopted to preserve primitive information. Then a multi-scale feature fusion module is applied to fuse multi-scale features from the decoder and information pool. Finally, the fused features are utilized to reconstruct the high-resolution MS image. Extensive experiments demonstrate that our proposed method achieves favorable performance against other methods in terms of quantitative metrics and visual quality. Besides, the results on running time indicate that our method can achieve real-time performance.  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
10  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:000619406500001  
Indexed Date:  
2021-03-10  
Institution Address:  
Sichuan Univ, Coll Elect & Informat Engn, Chengdu 610064, Sichuan, Peoples R China  
Institution Address:  
Xidian Univ, Sch Elect Engn, Xian 710071, Peoples R China  
Institution Address:  
Incheon Natl Univ, Dept Embedded Syst Engn, Incheon 22012, South Korea  
Institution Address:  
TAL Educ Grp, Beijing, Peoples R China  
Institution Address:  
Sichuan Univ, Inst Nucl Sci & Technol, Chengdu 610064, Peoples R China  
Institution Address:  
Sichuan Univ, Sch Aeronut & Astronaut, Chengdu 610064, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
TAL Educ Grp  
Institution:  
Institution:

### 757. Mobile-Deep Based PCB Image Segmentation Algorithm Research

摘要: Aiming at the problems of inaccurate edge segmentation, the hole phenomenon of segmenting large-scale targets, and the slow segmentation speed of printed circuit boards (PCB) in the image segmentation process, a PCB image segmentation model Mobile-Deep based on DeepLabv3+ semantic segmentation framework is proposed. Firstly, the DeepLabv3+ feature extraction network is replaced by the lightweight model MobileNetv2, which effectively reduces the number of model parameters; secondly, for the problem of positive and negative sample imbalance, a new loss function is composed of Focal Loss combined with Dice Loss to solve the category imbalance and improve the model discriminative ability; in addition, a more efficient atrous spatial pyramid pooling (E-ASPP) module is proposed. In addition, a more efficient E-ASPP module is proposed, and the Roberts crossover operator is chosen to sharpen the image edges to improve the model accuracy; finally, the network structure is redesigned to further improve the model accuracy by drawing on the multi-scale feature fusion approach. The experimental results show that the proposed segmentation algorithm achieves an average intersection ratio of 93.45%, a precision of 94.87%, a recall of 93.65%, and a balance score of 93.64% on the PCB test set, which is more accurate than the common segmentation algorithms Hrnetv2, UNet, PSPNet, and PCBSegClassNet, and the segmentation speed is faster.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001126959900016  
Indexed Date:  
2024-01-08  
Institution Address:  
Fujian Univ Technol, Sch Elect Elect Engn & Phys, Fuzhou 350118, Peoples R China  
Institution Address:  
Xiapu Power Supply Co, State Grid Fujian Power Supply Co LTD, Ningde 355100, Fujian, Peoples R China  
Institution:  
Institution:  
Xiapu Power Supply Co

### 758. Classification of crop pests based on multi-scale feature fusion

摘要: Crop diseases and insect pests are a serious natural disaster, which needs to be predicted and monitored in time to ensure the output of crops. Due to the wide variety of pests and the similar morphology of crops in the early stages of growth, it is difficult for agricultural workers to accurately identify various types of pests. Crop insects have brought huge challenges to the prevention and control of plant diseases and insect pests. In response to this problem, we propose a way of classification of crop pests based on multi-scale feature fusion(MFFNet) to accurately recognizes and classifies crop pests. First, the multi-scale feature extraction module (MFE) is designed by using dilated convolution to obtain the multi-scale feature map of the pest image. At the same time, extracted the deep feature information of the image by the feature extraction module (DFE). Finally, the features extracted separately by the multi-scale feature extraction module (MFE) and the feature extraction module (DFE) were fused thus achieving accurately classified and identified the crops insects by the way of end-to-end. Experiments show that our proposed method has obtained excellent classification performance on the dataset of 12 types of pests, its classification accuracy rate (ACC) reached 98.2%.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
在 BIOSIS Citation Index 中的被引频次:  
16  
Web of Science 核心合集中的 "被引频次":  
37  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
40  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:000754276800005  
Indexed Date:  
2022-03-02  
Institution Address:  
Guangxi Univ, Nanning 530007, Peoples R China  
Institution Address:  
Guangxi Key Lab Mfg Syst & Adv Mfg Technol, Nanning 530007, Peoples R China  
Institution:  
Institution:  
Guangxi Key Lab Mfg Syst & Adv Mfg Technol

### 759. Track Fastener Defect Detection Model Based on Improved YOLOv5s

摘要: Defect detection of track fasteners is a prerequisite for safe and reliable railroad operation. The traditional manual visual inspection method has been unable to meet the growing demand for railroad network inspection in China. To achieve the need for accurate, fast, and intelligent detection of rail fasteners, this paper proposes a rail fastener defect detection model based on improved YOLOv5s. Firstly, the convolutional block attention module (CBAM) is added to the Neck network of the YOLOv5s model to enhance the extraction of essential features by the model and suppress the information of minor features. Secondly, a weighted bidirectional feature pyramid network (BiFPN) is introduced to realize the multi-scale feature fusion of the model. Finally, the K-means++ algorithm is used to re-cluster the dataset to obtain the anchor box suitable for the fastener dataset and improve the positioning ability of the model. The experimental results show that the improved model achieves an average mean precision (mAP) of 97.4%, a detection speed of 27.3 FPS, and a model memory occupancy of 15.5 M. Compared with the existing target detection model, the improved model has the advantages of high detection accuracy, fast detection speed, and small model memory occupation, which can provide technical support for edge deployment of rail fastener defect detection.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
16  
被引频次合计:  
17  
eISSN:  
1424-8220  
入藏号:  
WOS:001071440700001  
PubMed ID:  
37514751  
Indexed Date:  
2023-10-08  
Institution Address:  
Lanzhou Jiaotong Univ, Sch Mech Engn, Lanzhou 730070, Peoples R China  
Institution Address:  
Tongji Univ, Shanghai Key Lab Rail Infrastruct Durabil & Syst S, Shanghai 201804, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, State Key Lab Tract Power, Chengdu 610031, Peoples R China  
Institution:  
Institution:  
Institution:

### 760. Cross attention guided multi-scale feature fusion for false-positive

摘要: False-positive reduction is a crucial step of computer-aided diagnosis (CAD) system for pulmonary nodules detection and it plays an important role in lung cancer diagnosis. In this paper, we propose a novel cross attention guided multi-scale feature fusion method for false-positive reduction in pulmonary nodule detection. Specifically, a 3D SENet50 fed with a candidate nodule cube is applied as the backbone to acquire multi-scale coarse features. Then, the coarse features are refined and fused by the multi-scale fusion part to achieve a better feature extraction result. Finally, a 3D spatial pyramid pooling module is used to enhance receptive field and a distributed aligned linear classifier is applied to get the confidence score. In addition, each of the five nodule cubes with different sizes centering on every testing nodule position is fed into the proposed framework to obtain a confidence score separately and a weighted fusion method is used to improve the generalization performance of the model. Extensive experiments are conducted to demonstrate the effectiveness of the classification performance of the proposed model. The data used in our work is from the LUNA16 pulmonary nodule detection challenge. In this data set, the number of true-positive pulmonary nodules is 1,557, while the number of false-positive ones is 753,418. The new method is evaluated on the LUNA16 dataset and achieves the score of the competitive performance metric (CPM) 84.8%.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在中国科学引文数据库中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
7  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:000906929200001  
PubMed ID:  
36401972  
Indexed Date:  
2023-01-19  
Institution Address:  
Jiangnan Univ, Jiangsu Prov Engn Lab Pattern Recognit & Computat, 1800 Lihu Ave, Wuxi 214122, Jiangsu, Peoples R China  
Institution Address:  
Jiangnan Univ, Coll Internet Things Engn, 1800 Lihu Ave, Wuxi 214122, Jiangsu, Peoples R China  
Institution Address:  
Jiangnan Univ, Dept Respirat, Affiliated Hosp, 1000 Hefeng Rd, Wuxi 214122, Jiangsu, Peoples R China  
Institution:  
Institution:  
Institution:

### 761. Adaptive local neighborhood search and dual attention convolution

摘要: In indoor scenes, due to the various semantic categories and scale differences of objects, the complex spatial structure of objects, and the different reflection characteristics of laser resulting in the uneven density distribution and the complex and diverse detail features of the point clouds, making the semantic segmentation of point clouds still a challenging task. In this paper, we propose an adaptive local neighborhood search and dual attention convolution network (ADNet) for complex semantic segmentation towards indoor point clouds. First, an adaptive local neighborhood search (ALNS) mechanism is constructed to address the problem that the traditional KNN is difficult to adapt to the changing density of complex point clouds. This mechanism dynamically adjusts the neighborhood radius according to the point cloud density, allowing for the construction of graphs that enable the network to learn local features with optimal receptive fields. Additionally, we propose a multi-scale feature fusion module to effectively integrate features of different scales by utilizing dynamic context-aware attention convolution (DCAA-Conv) to capture key context information between distant points and graph space attention convolution (GSA-Conv) to capture complex spatial geometric relations between nearby points and avoid the interference of noisy point clouds. Finally, spatial pyramid pooling specialized for indoor point clouds (SPP-IPCs) is employed for multi-scale feature fusion to enhance the ability to capture and express features at various scales. The proposed ADNet demonstrates proficiency in identifying minor categories and distinguishing building structures within complex environments, achieving superior segmentation accuracy compared with existing work on ScanNet v2 and S3DIS datasets, thus providing anew technical approach for accurate and complex 3D scene analysis.  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001414690600001  
Indexed Date:  
2025-02-10  
Institution Address:  
Xian Key Lab Image Proc Technol & Applicat Publ Se, Xian 710121, Shaanxi, Peoples R China  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Commun & Informat Engn, Xian 710121, Shaanxi, Peoples R China  
Institution Address:  
Int Joint Res Ctr Wireless Commun & Informat Proc, Xian 710121, Shaanxi, Peoples R China  
Institution Address:  
Shandong Univ, Sch Control Sci & Engn, Jinan 250061, Shandong, Peoples R China  
Institution:  
Xian Key Lab Image Proc Technol & Applicat Publ Se  
Institution:  
Institution:  
Int Joint Res Ctr Wireless Commun & Informat Proc  
Institution:

### 762. M<SUP>3</SUP>HOGAT: A Multi-View Multi-Modal Multi-Scale High-Order

摘要: Numerous scientific studies have found a link between diverse microorganisms in the human body and complex human diseases. Because traditional experimental approaches are time-consuming and expensive, using computational methods to identify microbes correlated with diseases is critical. In this paper, a new microbe-disease association prediction model is proposed that combines a multi-view multi-modal network and a multi-scale feature fusion mechanism, called M(3)HOGAT. Firstly, a microbe-disease association network and multiple similarity views are constructed based on multi-source information. Then, consider that neighbor information from disparate orders might be more adept at learning node representations. Consequently, the higher-order graph attention network (HOGAT) is devised to aggregate neighbor information from disparate orders to extract microbe and disease features from different networks and views. Given that the embedding features of microbe and disease from different views possess varying importance, a multi-scale feature fusion mechanism is employed to learn their interaction information, thereby generating the final feature of microbes and diseases. Finally, an inner product decoder is used to reconstruct the microbe-disease association matrix. Compared with five state-of-the-art methods on the HMDAD and Disbiome datasets, the results of 5-fold cross-validations show that M(3)HOGAT achieves the best performance. Furthermore, case studies on asthma and obesity confirm the effectiveness of M(3)HOGAT in identifying potential disease-related microbes.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2168-2194  
eISSN:  
2168-2208  
入藏号:  
WOS:001329782300013  
PubMed ID:  
39012741  
Indexed Date:  
2024-10-17  
Institution Address:  
Qufu Normal Univ, Sch Comp Sci, Rizhao 276826, Shandong, Peoples R China  
Institution Address:  
Univ Hlth & Rehabil Sci, Sch Hlth & Life Sci, Qingdao 266000, Shandong, Peoples R China  
Institution Address:  
Qufu Normal Univ, Qufu Normal Univ Lib, Rizhao 276826, Shandong, Peoples R China  
Institution:  
Institution:  
Institution:

### 763. Image registration combining cross-scale point matching and multi-scale

摘要: Image registration plays an important role in computer-aided diagnosis of brain diseases and remote surgery. The U-Net and its variants have been widely used in the field of medical image registration, , achieving good results in registration accuracy and time. However, , existing registration models have difficulty in learning the edge features of small structures in complex image deformations and ignore the correlation of contextual information at different scales. To address these issues, , a registration model is proposed based on cross-scale point matching combined with multi-scale feature fusion. Firstly, , a cross-scale point matching module is introduced into encoding structure of the model to enhance the representation of prominent region features and grasp the edge details of small structure features. Then, , multi-scale features are fused in the decoding structure to form a more comprehensive feature description. Finally, , an attention module is integrated into the multi-scale feature fusion module to highlight spatial and channel information. The experimental results on three brain Magnetic Resonance (MR ) datasets show that, , taking the OASIS-3 dataset as an example, , the registration accuracy has been improved by 23. 5%, , 12. 4%, , 0. 9%, , and 2. 1% compared to methods such as Affine, SyN, VoxelMorph and CycleMorph, , respectively. The corresponding ASD values for each method have decreased by 1. 074, 0. 434, 0. 043, and 0. 076. The proposed model can better grasp the feature information of images, , which improves registration accuracy and has important implications for the development of medical image registration.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1007-2780  
eISSN:  
2097-3217  
入藏号:  
WOS:001295070000009  
Indexed Date:  
2024-09-10  
Institution Address:  
Inner Mongolia Univ Sci & Technol, Sch Informat Engn, Baotou 014010, Peoples R China  
Institution Address:  
Inner Mongolia Univ Technol, Sch Informat Engn, Hohhot 010051, Peoples R China  
Institution:  
Institution:

### 764. CRNet: Cascaded Refinement Network for polyp segmentation

摘要: Technology for automatic segmentation plays a crucial role in the early diagnosis and treatment of ColoRectal Cancer (CRC). Existing polyp segmentation methods often focus on advanced feature extraction while neglecting detailed low-level features, This somewhat limits the enhancement of segmentation performance. This paper proposes anew technique called the Cascaded Refinement Network (CRNet), designed to improve polyp segmentation performance by combining low-level and high-level features through a cascaded contextual network structure. To accurately capture the morphological variations of polyps and enhance the clarity of segmentation boundaries, we have designed the Multi-Scale Feature Optimization (MFO) module and the Contextual Edge Guidance (CEG) module. Additionally, to further enhance feature fusion and utilization, we introduced the Cascaded Local Feature Fusion (CLFF) module, which effectively integrates cross-layer correlations, allowing the network to understand complex polyp structures better. By conducting a large number of experiments, our model achieved a 0.3% and 3.1% higher mDice score than the latest MMFILNet in the two main datasets of Kvasir-SEG and CVC-ClinicDB, respectively. Ablation studies show that MFO improves the baseline score by 4%, and the network without CLFF and CEG results in a reduction of 2.4% and 1.7% in mDice scores, respectively. This further validates the contribution of each module to the polyp segmentation performance. CRNet enhances model performance through the introduction of multiple modules but also increases model complexity. Future work will explore how to reduce computational complexity and improve inference speed while maintaining high performance. The source code for this paper can be found at https://github.com/l1986036/CRNet.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1319-1578  
eISSN:  
2213-1248  
入藏号:  
WOS:001399885500001  
Indexed Date:  
2025-01-23  
Institution Address:  
Guangxi Univ Sci & Technol, Sch Automat, Liuzhou 545006, Peoples R China  
Institution Address:  
Guangxi Univ Sci & Technol, Dept Med, Liuzhou 545006, Peoples R China  
Institution:  
Institution:

### 765. UAV small target detection algorithm based on an improved YOLOv5s model

摘要: The targets of UAV target detection are usually small targets, and the backgrounds are complex. In this work, aiming at the problem that small targets are easy to be missed or misdetected during the UAV detection, an improved YOLOv5s\_MSES target detection algorithm based on YOLOv5s is proposed. First of all, to solve the problem of UAV's difficulty in detecting small targets, the detection layer is ameliorated into the small target detection layer STD, which makes the model more easily detect the small targets. Then, the multi-scale feature fusion module is added to improve the detection accuracy of the small targets. Furthermore, by combining multi-scale module and attention module, a new connection method is proposed to retain the large scale of feature information. Finally, in contrast with some existent methods, the experimental results of VisDrone2019 UAV target detection dataset show that our proposed YOLOv5s\_MSES can achieve the better detection effect, and more effectively complete the small target detection task for UAV aerial photography images.  
Web of Science 核心合集中的 "被引频次":  
17  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
18  
ISSN:  
1047-3203  
eISSN:  
1095-9076  
入藏号:  
WOS:001085938900001  
Indexed Date:  
2023-11-02  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Sch Automat Engn, Nanjing 211106, Jiangsu, Peoples R China  
Institution Address:  
Nanjing Forestry Univ, Sch Informat Sci & Technol, Nanjing 210037, Jiangsu, Peoples R China  
Institution:  
Institution:

### 766. Multi-scale object detection algorithm combined with super-resolution

摘要: At present,most object detection algorithms have poor performance because of the large span of scales,leading to errors and omissions.To address the above issues,a multi-scale object detection algorithm combined with the super-resolution technology is proposed in this paper.First,based on the one-stage YOLO framework,the super-resolution module is employed to the neck network during the process of multi-scale feature fusion,which avoids further loss of detailed features in deeper layers.Second,the attention module is integrated in the shallower layers to focus on the channel information on object contour features and to suppress irrelevant features,thus improving the superficial representational capacity.Finally,ablation and comparative experiments are carried out on PASCAL VOC 2007and MS COCO 2017public datasets. Experimental results show that the proposed module can improve the detection performance.Compared with the current contrast algorithms,not only can the average accuracy rate of small,medium and large objects be increased by 1.20%,1.20%and 1.30%,but also the average recall rate can be improved by 4.20%,3.50% and 4.20%,respectively.  
摘要:  
目前大多数目标检测算法,由于尺度跨度较大而导致模型整体精确率和召回率不高,容易出现错检、漏检等现象。针对上述问题,提出一种融合超分辨率重建技术的多尺度目标检测算法。首先,算法以单阶段目标检测算法YOLO框架为基础,在颈部网络实现多尺度特征融合时加入超分辨率重建模块,避免进一步丢失较深层特征图中的细节特征。其次,使用通道注意力模块将较浅层特征图中的无关特征进行抑制,重点关注含有目标轮廓特征的通道信息,进一步增强浅层特征的表达能力。最后,在PASCAL VOC 2007和MS COCO 2017公开数据集上进行了消融实验和对比实验。实验结果表明,所提模块对检测性能有不同程度的提升,相比当前其他多尺度目标检测算法,所提算法在大、中、小三种尺度下目标平均精确率分别提升约1.20%、1.20%和1.30%,平均召回率分别提升约4.20%、3.50%和4.20%,算法整体检测性能得到进一步改善。  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
3  
ISSN:  
1001-2400  
入藏号:  
CSCD:7501215  
Indexed Date:  
2023-10-14  
Institution Address:  
湖北工业大学, 太阳能高效利用与储能运行控制湖北省实验室, 武汉, 湖北 430068, 中国  
Institution Address:  
Hubei University of Technology, Hubei Laboratory of Solar Energy Efficient Utilization and Energy Storage Operation Control, Wuhan, Hubei 430068, China  
Institution:  
湖北工业大学  
Institution:  
Hubei University of Technology

### 767. Intelligent Detection of Steel Defects Based on Improved Split Attention

摘要: The intelligent monitoring and diagnosis of steel defects plays an important role in improving steel quality, production efficiency, and associated smart manufacturing. The application of the bio-inspired algorithms to mechanical engineering problems is of great significance. The split attention network is an improvement of the residual network, and it is an improvement of the visual attention mechanism in the bionic algorithm. In this paper, based on the feature pyramid network and split attention network, the network is improved and optimised in terms of data enhancement, multi-scale feature fusion and network structure optimisation. The DF-ResNeSt50 network model is proposed, which introduces a simple modularized split attention block, which can improve the attention mechanism of cross-feature graph groups. Finally, experimental validation proves that the proposed network model has good performance and application prospects in the intelligent detection of steel defects.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
21  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
51  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
51  
ISSN:  
2296-4185  
入藏号:  
WOS:000748113100001  
PubMed ID:  
35096796  
Indexed Date:  
2022-02-06  
Institution Address:  
Wuhan Univ Sci & Technol, Key Lab Met Equipment & Control Technol, Minist Educ, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Hubei Key Lab Mech Transmiss & Mfg Engn, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Precis Mfg Res Inst, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Res Ctr Biomimet Robot & Intelligent Measurement, Wuhan, Peoples R China  
Institution Address:  
Three Gorges Univ, Hubei Key Lab Hydroelect Machinery Design & Maint, Yichang, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 768. Improved YOLOv7 Object Detection Algorithm for Fisheye Images

摘要: Images taken by fisheye cameras are characterized by wide field of view,geometric distortion and large scale variance,which bring great challenges to object detectors based on general convolutional networks.Existing object detection algorithms can be further improved with respect to network structure design,feature learning to be applicable to the distorted object detection task on fisheye images.To mitigate the effect of radial distortion on fisheye images,a multi-head attention module with multi-branch stacking structure is used in the YOLOv7 backbone to capture global contextual information.Meanwhile,a simple and efficient layer aggregation structure combining deformable convolutions is used on the Neck side of YOLOv7 to achieve effective multi-scale feature fusion.Experiments are conducted on the public comprehensive fisheye image dataset VOC\_360,and the results show that the improved YOLOv7 fisheye image object detector effectively achieves detection accuracy of 84.3% and 70.4% for mAP50 and mAP50:95,respectively,which is 3.1 percentage points and 6.4 percentage points higher than the baseline model YOLOv7,respectively.  
摘要:  
鱼眼相机捕获的图像具有宽视场、几何失真和尺度差异大等特点,这给基于标准卷积网络的目标检测器带来了巨大的挑战。现有的目标检测算法可以在网络结构设计、特征学习等方面进一步改进以适用于鱼眼图像上的失真目标检测任务。为减轻鱼眼图像上径向畸变的影响,研究在YOLOv7主干引入多分支堆叠结构的多头注意力模块以捕获全局上下文信息,提高检测准确性。同时,在YOLOv7的Neck侧,使用简单高效的融合可变形卷积的层聚合结构以实现有效的多尺度特征融合,提高模型对失真目标的特征提取能力。提出的检测模型直接在鱼眼图像上执行,无须指定先验信息和校准。在公开的综合鱼眼图像数据集VOC\_360上进行实验,结果表明,改进后的YOLOv7鱼眼图像目标检测器有效地提高了检测精度,mAP50、mAP50:95分别达到84.3%、70.4%,相比基准模型YOLOv7分别提升3.1个百分点、6.4个百分点。  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7794600  
Indexed Date:  
2024-11-01  
Institution Address:  
北京联合大学, 北京市信息服务工程重点实验室, 北京 100101, 中国  
Institution Address:  
北京理工大学计算机学院, 北京 100081, 中国  
Institution Address:  
山东财经大学计算机科学与技术学院, 济南, 山东 250014, 中国  
Institution Address:  
Beijing Union University, Beijing Key Laboratory of Information Service Engineering, Beijing 100101, China  
Institution Address:  
School of Computer Science and Technology,Beijing Institute of Technology, Beijing 100081, China  
Institution Address:  
School of Computer Science and Technology,Shandong University of Finance and Economics, Jinan, Shandong 250014, China  
Institution:  
北京联合大学  
Institution:  
北京理工大学计算机学院  
Institution:  
山东财经大学计算机科学与技术学院  
Institution:  
Beijing Union University  
Institution:  
School of Computer Science and Technology,Beijing Institute of Technology  
Institution:  
School of Computer Science and Technology,Shandong University of Finance and Economics

### 769. Cross-PIC: A cross-scale in-context learning network for 3D multibeam

摘要: The multibeam sonar has become a powerful tool for undersea pipeline detection, which can obtain high quality three-dimensional point cloud data. However, the identification accuracy of current pipeline algorithms remains low because of insufficient utilization of context features. Therefore, this paper proposes an improved context-learning point cloud segmentation network, named Cross-scale Point-In-Context (Cross-PIC), to identify submarine pipelines. Cross-PIC designs the sampling module Balanced-FPS to improve the equalization sampling. Meanwhile, Cross-PIC constructs different scale coding architecture based on context learning, which includes Auxiliary line and Baseline for multi-scale feature encoding. Finally, the Refine Module is designed to realize cross-scale coded feature fusion to output segmentation results. Experiments with different fuzzy degrees show that Cross-PIC also significantly outperforms the original Point-In-Context network with less training data. Cross-PIC also achieves better results than other networks based on point, voxel, and dynamic graph convolution.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
0029-8018  
eISSN:  
1873-5258  
入藏号:  
WOS:001367429300001  
Indexed Date:  
2024-12-07  
Institution Address:  
China Univ Petr East China, Coll Oceanog & Space Informat, Qingdao 266580, Peoples R China  
Institution Address:  
Minist Nat Resources, Technol Innovat Ctr Maritime Silk Rd Marine Resour, Qingdao, Peoples R China  
Institution Address:  
China Univ Petr East China, Coll Comp Sci & Technol, Qingdao 266580, Peoples R China  
Institution:  
Institution:  
Institution:

### 770. Real-Time ConstructionWorker Detection Method for Edge Device

摘要: Construction worker detection has an important application value in construction management. The complex images backgrounds and diverse perspectives at construction sites make the task of construction worker detection difficult, while most of the construction sites have poor infrastructure facilities and poor network conditions, making them unsuitable for model deployment on large GPU workstations. To address the above issues, it proposes a real-time worker detection algorithm for construction scenes based on YOLOv3 detection network, adding a feature pyramid pooling module, increasing multi-scale feature fusion and improving candidate frames to improve detection accuracy, and using a channel pruning algorithm to lighten the detection network to adapt to the computing power of edge- end devices. This method achieves an average accuracy of 88.23% on the homemade construction worker dataset, which is 4.89 percentage points higher than the YOLOv3 detection method, and compresses the model size to 1/13 of the original one, doubles the detection speed. It achieves a detection speed of 69.08 frame/s on the embedded end device, meeting the requirements for realtime edge-end detection at construction sites.  
摘要:  
施工人员检测在施工管理工作中有重要的应用价值。施工现场图像背景复杂且视角多样,给施工人员检测任务带来难度,同时施工现场大多基础配套设施不完善,并且网络条件较差,不适合在大型GPU工作站上进行模型部署。针对以上问题,以YOLOv3检测网络为基础,加入特征金字塔池化模块,增加多尺度特征融合并改进候选框,提升检测精度,同时采用通道剪枝算法对检测网络进行轻量化处理以适应边缘端设备算力,提出一种面向边缘端的施工人员实时检测方法。该方法在自制的施工人员数据集上平均准确率可达到88.23%,较YOLOv3检测方法提升4.89个百分点,且将模型大小压缩至原来的1/13,检测速度提升一倍,在嵌入式设备Jetson Xavier NX上检测速度可达到69.08帧/s,满足在施工现场进行实时边缘端检测的要求。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1002-8331  
入藏号:  
CSCD:7104753  
Indexed Date:  
2022-02-11  
Institution Address:  
中国海洋大学工程学院, 青岛, 山东 266100, 中国  
Institution Address:  
School of Engineering, Ocean University of China, Qingdao, Shandong 266100, China  
Institution:  
中国海洋大学工程学院  
Institution:  
School of Engineering, Ocean University of China

### 771. Segmentation of low scattering region in SAR images using multi-module

摘要: The proposed multi-module fusion network (MMFNet) is designed for the segmentation of low scattering regions such as roads, waters, and shadows in synthetic aperture radar (SAR) images in this paper. It is primarily comprised of three modules, i.e. high-resolution backbone network module, spatial pyramid pooling convolution (SPPC) module, and channel attention module, and trained with weighted cross-entropy loss. The high-resolution backbone network works to retain high resolution of feature maps and reduce spatial accuracy loss, which contributes to the extraction of edge information. SPPC module performs multi-scale feature fusion, extracts target areas with different sizes and improves network accuracy. Channel attention module intensifies network expression of category information, thus further improves network performance. Our experimental analysis using real SAR data shows that MMFNet achieves good low scattering region segmentation, with mean IoU (MIoU) reaching up to 82.5%.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:000874329900001  
Indexed Date:  
2022-11-04  
Institution Address:  
Univ Elect Sci & Technol China, Sch Resources & Environm, 2006 Xiyuan Ave, Chengdu 611731, Sichuan, Peoples R China  
Institution Address:  
Sichuan Aerosp Elect Equipment Res Inst, Chengdu, Peoples R China  
Institution Address:  
Univ Elect Sci & Technol China, Sch Informat & Commun Engn, Chengdu, Peoples R China  
Institution:  
Institution:  
Sichuan Aerosp Elect Equipment Res Inst  
Institution:

### 772. MSPTS-Net: Multiscale Parallel TemporalSpatial Network for Precipitation

摘要: Precipitation nowcasting plays an important role in people's daily lives, yet current prediction methods still encounter some challenges. Most radar echo extrapolation methods rely on a single-input-single-output (SISO) structure, which can accumulate errors over time, resulting in a decline in prediction accuracy for longer forecasts. Furthermore, the radar echoes exhibit complex characteristics, making it difficult to capture underlying dynamics solely through the extraction of features at a single scale, especially for high-intensity rainfall events. To overcome these difficulties, this letter proposes a multiscale parallel temporal-spatial network (MSPTS-Net) based on a multi-input-multi-output (MIMO) architecture, which reduces the accumulation error for longer predictions. Specifically, we employ a multiscale temporal feature extraction module to capture both long-term and short-term temporal evolution patterns. In the spatial dimension, we adopt a global-to-local multiscale feature extraction module to model and represent the characteristics of radar echoes across different scales. For better representation of temporal-spatial characteristics, the self-attention mechanism is employed to integrate multiscale features across both temporal and spatial dimensions. During the training process, we use a hybrid loss function by combining mean squared error (mse) and Charbonnier loss to enhance prediction accuracy. Compared with existing methods, the proposed MSPTS-Net structure demonstrates significant advantages in performance.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001398604700007  
Indexed Date:  
2025-01-25  
Institution Address:  
Shanghai Univ, Sch Commun & Informat Engn, Shanghai Inst Adv Commun & Data Sci, Shanghai 200444, Peoples R China  
Institution:

### 773. Development of a cross-scale weighted feature fusion network for

摘要: Surface defects of hot-rolled steel would affect the performance and appearance of the final products. In order to detect steel surface defects efficiently, a cross-scale weighted feature fusion network for identifying defect categories and locating defects is proposed in this work. Combined with Laplace sharpening, the backbone in the YOLOv5s model is used to extract multi-scale defect features from input images. And then, an improved weighted bi-directional feature pyramid network embedded with residual modules is proposed to aggregate multi-scale feature maps for enhancing the robustness of multi-size defect representation. Finally, four prediction branches accompanied with prior bounding boxes by a k-means clustering algorithm are responsible for predicting defects with different sizes. The proposed detection network is verified on the NEU-DET dataset, and experimental results show that the proposed network can achieve 86.8% mAP with the IoU threshold of 0.5, and can efficiently process images at 51 fps with the RGB image size 640 x 640. The Laplace sharpening module, the k\_means clustering module and the improved C3-BiFPN module all contribute to the improvement of performance (mAP) of the proposed network by 1.8%, 2.7% and 3.8%, respectively. Our experimental results demonstrate that the proposed framework can effectively detect the surface defects of hot-rolled steel, and has potential to be used for real-time surface defect detection. Meanwhile, the versatility of the proposed network for other types of defect detection is also evaluated on the MT dataset and the DAGM dataset.  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
45  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
47  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:000917296700013  
Indexed Date:  
2023-02-09  
Institution Address:  
Hefei Univ Technol, Sch Instrument Sci & Optoelect Engn, Hefei 230009, Anhui, Peoples R China  
Institution Address:  
Hefei Univ Technol, Anhui Prov Key Lab Measuring Theory & Precis Instr, Hefei 230009, Anhui, Peoples R China  
Institution:  
Institution:

### 774. Apple Leaf Disease Recognition and Sub-Class Categorization Based on

摘要: Apple diseases cause a lot of economic losses to fruit growers in China. Early diagnosis and accurate recognition of apple diseases can control the spread of disease and reduce production costs. However, the significance of disease characteristic of apple leaves in complex environment is relatively weak, and the fine-grain among different diseases of apple leaves is high, and the conventional feature extraction methods will lose the discrimination information. To solve these problems, an apple disease classification model based on multi-scale feature fusion is proposed in this paper. Firstly, the information flow of conventional residual network (ResNet) was improved to achieve efficient information circulation through changing the position of batch normalization and rectified linear unit (ReLU). Secondly, in order to solve the problem of serious loss of information in ResNet downsample, the channel projection and spatial projection of downsample were separated. Lastly, the 3 x 3 conv in ResBlocks was replaced by pyramid convolution, and the dilated convolution with different dilation rate was introduced into pyramid convolution to enhance the output scale of feature maps and improve the robustness of the model. The optimized model was verified on the dataset of this paper, and the optimized model had stronger anti-noise ability and better robustness, excellent learning effect and fast convergence speed. The classification accuracy on the original dataset is 94.24%, and that on the preprocessed dataset is 94.99%. The results demonstrate that the optimal model has a high accuracy, which can provide a reference for the prevention and control of apple leaf diseases.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
25  
在中国科学引文数据库中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
6  
被引频次合计:  
26  
ISSN:  
2169-3536  
入藏号:  
WOS:000673660300001  
Indexed Date:  
2021-07-25  
Institution Address:  
Jiangsu Univ, Sch Elect & Informat Engn, Zhenjiang 212013, Jiangsu, Peoples R China  
Institution:

### 775. DS-YOLO: A dense small object detection algorithm based on inverted

摘要: In the field of security, intelligent surveillance tasks often involve a large number of dense and small objects, with severe occlusion between them, making detection particularly challenging. To address this significant challenge, Dense and Small YOLO (DS-YOLO), a dense small object detection algorithm based on YOLOv8s, is proposed in this paper. Firstly, to enhance the dense small objects' feature extraction capability of backbone network, the paper proposes a lightweight backbone. The improved C2fUIB is employed to create a lightweight model and expand the receptive field, enabling the capture of richer contextual information and reducing the impact of occlusion on detection accuracy. Secondly, to enhance the feature fusion capability of model, a multi-scale feature fusion network, Light-weight Full Scale PAFPN (LFS-PAFPN), combined with the DO-C2f module, is introduced. The new module successfully reduces the miss rate of dense small objects while ensuring the accuracy of detecting large objects. Finally, to minimize feature loss of dense objects during network transmission, a dynamic upsampling module, DySample, is implemented. DS-YOLO was trained and tested on the CrowdHuman and VisDrone2019 datasets, which contain a large number of densely populated pedestrians, vehicles and other objects. Experimental evaluations demonstrated that DS-YOLO has advantages in dense small object detection tasks. Compared with YOLOv8s, the Recall and mAP@0.5 are increased by 4.9% and 4.2% on CrowdHuman dataset, 4.6% and 5% on VisDrone2019, respectively. Simultaneously, DSYOLO does not introduce a substantial amount of computing overhead, maintaining low hardware requirements. (c) 2024 The Author(s). Published by Elsevier B.V. on behalf of Shandong University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2097-0242  
eISSN:  
2667-3797  
入藏号:  
WOS:001349404000001  
Indexed Date:  
2024-11-15  
Institution Address:  
Qilu Univ Technol, Sch Informat & Automat Engn, Jinan 250353, Peoples R China  
Institution Address:  
Unimat Intelligent Technol Co, Jinan 250101, Peoples R China  
Institution Address:  
Univ St Andrews, Sch Comp Sci, St Andrews KY16 9AJ, Scotland  
Institution:  
Institution:  
Unimat Intelligent Technol Co  
Institution:

### 776. A lightweight dense crowd density estimation network for efficient

摘要: Crowd density estimation is a task of intelligent applications, and its operation efficiency is very important. However, to obtain a better density estimation performance, most of the existing works often design larger and more complex network structures, which will result in them occupying considerable memory, time and other resources at runtime, and require the support of high-performance hardware platforms, which are difficult to apply in practice. In this paper, to overcome the above problems, we propose a lightweight dense crowd estimation method based on channel attention multi-scale feature fusion. Specifically, in the process of feature extraction, an efficient and lightweight convolution module (L-weight) is designed to extract crowd features in stages, which reduces the amount of network parameters and computing costs, and we capture multi-scale crowd information through the feature extraction network of pyramid structure, which solves the problem of uneven crowd scale in video images. In the process of feature fusion, a channel attention fusion module is designed, which weights and fuses the feature information of different scales, effectively fuses multi-scale information and suppresses useless information. In addition, we design a new loss function, which enhances the sensitivity of the crowd through the pixel space loss (L2), counting loss (LC) and structural similarity loss (LS), to ensure the counting accuracy. Extensive experiments on four mainstream datasets demonstrate that compared with other state-of-the-art methods, our method achieves an optimal trade-off between counting performance and running speed, and is suitable for low-performance computing platforms such as embedded.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001097384700001  
Indexed Date:  
2023-11-29  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao 266590, Peoples R China  
Institution:

### 777. Lightweight method of feature point extraction and matching

摘要: To address the issues of the feature-matching method based on the SuperPoint network, such as low accuracy in feature-point extraction and high computational cost under challenges of lighting, pose and angles, a lightweight feature point extraction and matching method under a progressive strategy is put forward. Firstly, to reduce the model's computational cost, the SuperPoint network is modified using depthwise separable convolution. Secondly, an attention module is built in the feature extraction part to strengthen the network's spatial feature extraction capability. Also, a progressive multi-scale feature fusion module is designed to capture object details and boost feature representation capabilities. Finally, the obtained feature points are matched using the SuperGlue algorithm. Experimental analysis on the Hpatches dataset shows that the proposed algorithm achieves an average matching accuracy (mAP ) of 86 degrees o and feature point repeatability (Rep ) of 70 degrees o in illumination change scenarios, and mAP of 78 degrees o and Rep of 68 degrees o in viewpoint change scenarios. The proposed algorithm not only shows certain advantages in feature matching, but also achieves good results when applied to video stitching.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1007-2780  
eISSN:  
2097-3217  
入藏号:  
WOS:001351944100012  
Indexed Date:  
2024-11-17  
Institution Address:  
Taiyuan Normal Univ, Sch Comp Sci & Technol, Jinzhong 030619, Peoples R China  
Institution Address:  
Taiyuan Normal Univ, Shanxi Prov Key Lab Intelligent Optimizat Comp & B, Jinzhong 030619, Peoples R China  
Institution Address:  
Northwest Normal Univ, Sch Phys & Elect Engn, Lanzhou 730070, Peoples R China  
Institution:  
Institution:  
Institution:

### 778. Enhancing Rooftop Photovoltaic Segmentation Using Spatial Feature

摘要: Amidst the dual challenges of energy shortages and global warming, photovoltaic (PV) power generation has emerged as a critical technology due to its efficient utilization of solar energy. Rooftops, as underutilized spaces, are ideal locations for installing solar panels, avoiding the need for additional land. However, the accurate and generalized segmentation of large-scale PV panel images remains a technical challenge, primarily due to varying image resolutions, large image scales, and the significant imbalance between foreground and background categories. To address these challenges, this paper proposes a novel model based on the Res2Net architecture, an enhanced version of the classic ResNet optimized for multi-scale feature extraction. The model integrates Spatial Feature Reconstruction and multi-scale feature aggregation modules, enabling effective extraction of multi-scale data features and precise reconstruction of spatial features. These improvements are particularly designed to handle the small proportion of PV panels in images, effectively distinguishing target features from redundant ones and improving recognition accuracy. Comparative experiments conducted on a publicly available rooftop PV dataset demonstrate that the proposed method achieves superior performance compared to mainstream techniques, showcasing its effectiveness in precise PV panel segmentation.  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
1996-1073  
入藏号:  
WOS:001393560800001  
Indexed Date:  
2025-01-15  
Institution Address:  
Chongqing Water Resources & Elect Engn Coll, Sch Elect Power Engn, Chongqing 404155, Peoples R China  
Institution Address:  
Chongqing Univ Technol, Sch Elect & Elect Engn, Chongqing 404155, Peoples R China  
Institution:  
Chongqing Water Resources & Elect Engn Coll  
Institution:

### 779. FM-YOLOv7: an improved detection method for mine personnel helmet

摘要: Due to the complex underground environment and the small object of the helmet, the detection accuracy is low when the original YOLOv7 algorithm is used to detect whether the mine personnel wears the helmet, which cannot be applied to the actual operation site. In response to this problem, we proposed an FM-YOLOv7 mine personnel helmet detection. First, to improve the feature extraction ability of the shallow network and enhance the representation ability of the model on the helmet, we propose the fused-MBCA (fused-MBConv with the coordinate attention) module. Second, to improve the detection ability of small objects, enable the fused features to obtain high-level semantic information and low-level details from different scales, and have more extensive receptive fields, we propose the multi-scale feature fusion efficient layer aggregation networks. Finally, to accelerate the convergence of the model and improve the regression accuracy, we use efficient intersection over union as the bounding box regression loss function. These experiments are based on the self-built mine personnel safety helmet dataset. The results show that the FM-YOLOv7 model outperformed the other six algorithms. The mAP@0.5 of the proposed model can reach 85.7%, which is 1.4% higher than the original YOLOv7 model. Also, the improved YOLOv7 model achieves 91 frames per second in detection speed, which detects whether the mine personnel wears a safety helmet in real time. (C) 2023 SPIE and IS&T  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:001060595700020  
Indexed Date:  
2023-09-22  
Institution Address:  
Xian Univ Sci & Technol, Sch Elect & Control Engn, Xian, Peoples R China  
Institution:

### 780. Research on wind speed behavior prediction method based on multi-feature

摘要: Wind power is just in place for a major improvement to the power grid in security and economy. More accurate forecast methods are emerged to benefit grid operators and their customers. A wind speed behavior prediction method based on multi-feature and multi-scale integrated learning (MFMS) is developed in this paper. The behavior characteristics are proposed to solve the problem of the temporal and spatial feature extraction distinctly of wind speed. Multi-scale feature fusion and spatial pyramid pooling algorithms are proposed to reduce the loss of micro-space-scale and short-time-scale information. Meanwhile, environmental features and background features are applied to consider the influence of the overall wind farm environment on the wind speed of local wind turbines. The wind speed data of sixteen wind turbines in a wind farm at Zhangjiakou in North China are used to verify the method. Results show that MAPE, RMSE and R2 of 4-h ultra-short-term wind speed prediction are 6.164%, 0.275 and 0.966, respectively, which are improved by 1.786%-6.757%, 0.051-0.216 and 0.014-0.074 compared with other methods. The method proposed in this paper can be applied to wind speed prediction and provide a guideline for wind farm power prediction.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
11  
被引频次合计:  
11  
ISSN:  
0360-5442  
eISSN:  
1873-6785  
入藏号:  
WOS:000868978700007  
Indexed Date:  
2022-11-02  
Institution Address:  
North China Elect Power Univ Baoding, Dept Power Engn, Baoding, Peoples R China  
Institution Address:  
North China Elect Power Univ Baoding, Dept Elect & Commun Engn, Baoding, Peoples R China  
Institution Address:  
North China Elect Power Univ Baoding, Dept Mech Engn, Baoding, Peoples R China  
Institution:  
Institution:  
Institution:

### 781. Infrared Vehicle Detection Algorithm Based on Improved Shuffle-RetinaNet

摘要: In view of the low detection accuracy and high complexity of current multi-scale vehicle detection algorithms in infrared scenes, an infrared vehicle detection algorithm based on Shuffle-RetinaNet is proposed. On the basis of RetinaNet, the algorithm uses ShuffleNetV2 as the feature extraction network. A dual-branch attention module channel attention module is proposed, which adopts the dual-branch structure and adaptive fusion and enhances the ability to extract the key features of the target in infrared images. To optimize the feature fusion, the algorithm integrates cross-scale connection and fast normalized fusion in some feature layers to enhance the multi-scale feature expression. The calibration factor is set to enhance the task interaction of classification and regression, and the accuracy of target classification and locating is increased. A series of experiments are conducted on a self-built infrared vehicle dataset to verify the effectiveness of the proposed algorithm. The detection accuracy of this algorithm for the self-built vehicle dataset is 92.9%, the number of parameters is 11.74x10(6), and the number of floating-point operations is 24.35x10(9). The algorithm exhibits better detection performance on the public dataset FLIR ADAS. Experimental results indicate that the algorithm has advantages in detection accuracy and model complexity, giving it good application value in multi-scale vehicle detection tasks in infrared scenes.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001180502700010  
Indexed Date:  
2024-03-20  
Institution Address:  
Tianjin Univ, Sch Microelect, Tianjin 300072, Peoples R China  
Institution:

### 782. Multi-scale context-guided feature elimination for ancient tower image

摘要: A multi-scale context-guided feature elimination classification method was proposed, for resolving the problems of ambiguous discriminative feature localization and complex scene interference in the classification task of ancient tower building images. First, a feature extraction network with MogaNet as the core was constructed, and multi-scale feature fusion was combined to fully explore the image information. Next, a context information extractor was designed to utilize the semantic context of the network to align and filter more discriminative local features, enhancing the ability to capture detailed features. Then, a feature elimination strategy was proposed to suppress fuzzy class features and background noise interference, and a loss function was designed to constrain fuzzy feature elimination and classification prediction. At last, a Chinese ancient tower architecture image dataset was established to provide data to support research on complex backgrounds and fuzzy boundaries in the field of finegrained image categorization. This method achieved 96.3% accuracy on the self-constructed ancient tower architecture dataset, and 92.4%, 95.3% and 94.6% accuracy on three fine-grained datasets, namely, CUB-200-2011, Stanford Cars and FGVC-Aircraft, respectively. The proposed method outperforms other comparison algorithms and enables accurate classification of images of ancient tower buildings.  
摘要:  
针对古塔建筑图像分类任务中难以准确定位判别性特征以及复杂场景干扰的问题,提出多尺度上下文引导特征消除的分类方法.构建以MogaNet为核心的特征提取网络,结合多尺度的特征融合以充分挖掘图像信息;设计上下文信息提取器,利用网络的语义上下文来对齐和过滤更具判别性的局部特征,加强网络捕捉细节特征的能力;提出特征消除策略,抑制模糊类特征和背景噪声干扰,并设计损失函数来约束模糊类特征消除和分类预测;建立中国古塔建筑图像数据集,为细粒度图像分类领域内针对复杂背景和模糊边界的研究提供数据支撑.实验结果表明,所提方法在自建的古塔建筑数据集上达到了96.3%的准确率,并在CUB-200-2011、Stanford Cars和FGVC-Aircraft这3个细粒度数据集上分别达到了92.4%、95.3%和94.6%的准确率,优于其他对比算法,可以实现古塔建筑图像的精确分类.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1008-973X  
入藏号:  
CSCD:7848665  
Indexed Date:  
2025-01-25  
Institution Address:  
西安建筑科技大学信息与控制工程学院;;建筑机器人陕西省高等学校重点实验室, ;;建筑机器人陕西省高等学校重点实验室, 西安;;, 陕西;; 710300;;, 中国  
Institution Address:  
College of Information and Control Engineering,Xi'an University of Architecture and Technology;;Key Laboratory of Construction Robots for Higher Education in Shaanxi Province, ;;Key Laboratory of Construction Robots for Higher Education in Shaanxi Province, Xi'an;;, ;; 710300;;  
Institution:  
西安建筑科技大学信息与控制工程学院;;建筑机器人陕西省高等学校重点实验室  
Institution:  
College of Information and Control Engineering,Xi'an University of Architecture and Technology;;Key Laboratory of Construction Robots for Higher Education in Shaanxi Province

### 783. Attention-based Graph Clustering Network with Dual Information

摘要: The field of attributed graph clustering has garnered increasing attention, particularly with the advent of graph convolutional network (GCN), which have deepened our understanding of learning both attribute and structural information in graphs. Existing graph deep embedding clustering methods typically learn attribute or structural information alone, or integrate attribute information into a learning network for structural information. However, these methods fail to fully integrate the available information. Therefore, we propose a novel deep attributed graph clustering method named A ttention-based Graph Clustering N etwork with D ual I nformation I nteraction (ADIIN). Specifically, an attention-based interaction fusion module is presented to adaptively incorporate two types of information and propagate the fused information to both networks interactively. Additionally, it can adjustively integrate information from each hidden layer at different scales based on attentional mechanisms. Furthermore, we design amore robust quadruple joint self-supervision strategy to align node attribute representation, linear fusion representation, and multi-scale feature fusion representation, thereby enhancing the clustering performance of the entire model. Extensive experiments conducted on several benchmark datasets demonstrate that our proposed method outperforms state-of-the-art deep clustering methods. Our code is publicly available at https://github.com/sliboo/ADIIN.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0950-7051  
eISSN:  
1872-7409  
入藏号:  
WOS:001407615700001  
Indexed Date:  
2025-02-05  
Institution Address:  
Beijing Univ Technol, Coll Comp Sci, Beijing, Peoples R China  
Institution Address:  
Beijing Jiaotong Univ, Sch Comp & Informat Technol, Beijing Key Lab Traff Data Anal & Min, Beijing, Peoples R China  
Institution:  
Institution:

### 784. UAV image object detection based on self-attention guidance and global

摘要: Unmanned aerial vehicle (UAV) image object detection has garnered considerable attentions in fields such as Intelligent transportation, urban management and agricultural monitoring. However, it suffers from key challenges of the deficiency in multi-scale feature extraction and the inaccuracy when processing complex scenes and small-sized targets in practical applications. To address this challenge, we propose a novel UAV image object detection network based on self-attention guidance and global feature fusion, named SGGF-Net. First, in order to optimizing feature extraction in global perspective and enhancing target localization precision, the global feature extraction module (GFEM) is introduced by exploiting the self-attention mechanism to capture and integrate long-range dependencies within images. Second, a normal distribution-based prior assigner (NDPA) is developed by measuring the resemblance between ground truth and the priors, which improves the precision of target position matching and thus handle the problem of inaccurate localization of small targets. Furthermore, we design an attention-guided ROI pooling module (ARPM) via a deep fusion strategy of multilevel features for optimizing the integration of multi-scale features and improving the quality of feature representation. Finally, experimental results demonstrate the effectiveness of the proposed SGGF-Net approach.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0262-8856  
eISSN:  
1872-8138  
入藏号:  
WOS:001317456200001  
Indexed Date:  
2024-09-27  
Institution Address:  
Shijiazhuang Tiedao Univ, Sch Informat Sci & Technol, Shijiazhuang 050043, Hebei, Peoples R China  
Institution Address:  
Hebei Key Lab Electromagnet Environm Effects & Inf, Shijiazhuang 050043, Hebei, Peoples R China  
Institution Address:  
Shijiazhuang Key Lab Artificial Intelligence, Shijiazhuang 050043, Hebei, Peoples R China  
Institution:  
Institution:  
Hebei Key Lab Electromagnet Environm Effects & Inf  
Institution:  
Shijiazhuang Key Lab Artificial Intelligence

### 785. Small Object Detection Algorithm Based on ATO-YOLO

摘要: Small object detection is of great significance in the field of computer vision. However, existing methods often suffer from issues such as missed detection and false alarms when dealing with challenges like scale variation, dense object arrangement, and irregular layouts. To address these problems, ATO-YOLO, an improved version of the YOLOv5 algorithm is proposed. Firstly, this paper introduces an adaptive feature extraction (AFE)module that incorporates an attention mechanism to enhance the feature representation capability of the detection model. By dynamically adjusting the weight allocation to highlight key object features, AFE improves the accuracy and robustness of object detection tasks in various scenarios. Secondly, a triple feature fusion (TFF)mechanism is designed to effectively utilize multi-scale information by fusing feature maps from different scales, resulting in more comprehensive object features and enhanced detection performance for small objects. Lastly, an output reconstruction (ORS)module is introduced, which removes the large object detection layer and adds a small object detection layer, enabling precise localization and recognition of small objects. This module also reduces model complexity and improves detection speed compared to the original model. Experimental results demonstrate that the ATO-YOLO algorithm achieves an mAP@0.5 of 38.2% on the VisDrone dataset, a 6.1 percentage points improvement over YOLOv5, with a relative FPS increase of 4.4%. This algorithm enables fast and accurate detection of small objects.  
摘要:  
小目标检测在计算机视觉领域具有重要意义,但现有方法在应对小目标的尺度变化、目标密集和无规则排列等挑战时经常出现漏检和误检的问题。为解决这些问题,提出基于改进YOLOv5算法的ATO-YOLO。为提升检测模型的特征表达能力,提出一种结合注意力机制的自适应特征提取模块(adaptive feature extraction,AFE),通过动态调整权重分配突出关键目标的特征表示,提高目标检测任务在不同场景下的准确性和鲁棒性。设计一种三重特征融合机制(triple feature fusion,TFF),能够在不同尺度下充分利用多尺度信息,将多个尺度的特征图融合,以获取更全面的目标特征,提升对小目标的检测效果。引入一种输出重构模块(output reconstruction,ORS),通过去除大目标检测层并增加小目标检测层,实现精确定位和识别小目标,并且相对于原模型复杂度更低,检测速度更快。实验结果表明,ATO-YOLO算法在VisDrone数据集上的mAP@0.5达到了38.2%,较原YOLOv5提升了6.1个百分点,且FPS较改进前提升了4.4%,能够快速准确地对小目标进行检测。  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1002-8331  
入藏号:  
CSCD:7694525  
Indexed Date:  
2024-06-21  
Institution Address:  
河北科技大学信息科学与工程学院, 石家庄, 河北 050018, 中国  
Institution Address:  
College of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang, Hebei 050018, China  
Institution:  
河北科技大学信息科学与工程学院  
Institution:  
College of Information Science and Engineering, Hebei University of Science and Technology

### 786. Anchor constrained refinement network with Intersection-over-Union-aware

摘要: Object detection has achieved good progress in the last few years. However, there are many challenges in the field of remote sensing imagery. Objects in remote sensing images usually have arbitrary orientations and various scales. In addition, some objects are easily overwhelmed by a cluttered background. To take advantage of single-stage object detectors that have fast speed, many cascaded structures based on single-stage detectors have been proposed to improve detection performance. However, feature inconsistency in cascade structure results in poor detection performance. To address these problems, we propose an innovative model in terms of both model improvement and loss function refinement. This model consists of an attention module to highlight useful information in cluttered scenes, a multi-scale feature fusion module, and a cascade refinement module with anchor constrained convolution to address feature inconsistency. Furthermore, Intersection-over-Union (IoU) classification loss is proposed to enhance the correlation between classification and localization, and a scale-aware regression loss is proposed to improve the detection performance on objects with different scales. We conducted extensive experiments on both the DOTA dataset and the HRSC2016 dataset, and the experimental results show that our model has advantages compared with current state-of-the-art methods. (C) 2022 SPIE and IS&T  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:000762340700033  
Indexed Date:  
2022-03-29  
Institution Address:  
Jiangnan Univ, Dept Comp Sci, Wuxi, Jiangsu, Peoples R China  
Institution Address:  
Jiangnan Univ, Jiangsu Prov Engn Lab Pattern Recognit & Computat, Wuxi, Jiangsu, Peoples R China  
Institution:  
Institution:

### 787. LGA-YOLO for Vehicle Detection in Remote Sensing Images

摘要: In remote sensing images, vehicles often appear on a minuscule scale, lacking features and easily overwhelmed by intricate background information. This becomes even more challenging in low illumination or occluded environments, leading to missed detections and false alarms. A novel vehicle detection algorithm, known as local and global aware YOLO (LGA-YOLO), is introduced to tackle these issues. LGA-YOLO incorporates two innovative and plug-and-play modules: the multiscale large kernel local aware module (MLKM) and the directional global context aware module (DGAM). MLKM widens the receptive field and enhances local features, while DGAM gathers global context information, highlighting vehicle features against complex backgrounds. Based on these modules, a high-low feature fusion network is reconstructed, capturing multiscale object features and effectively leveraging shallow features. Our self-constructed dataset (USOD), VEDAI, and DOTA are employed to validate LGA-YOLO's efficacy. In USOD, the results demonstrate the remarkable performance of LGA-YOLO, with precision, recall, AP(0.5), and AP(0.5:0.95) scores of 0.927, 0.889, 0.930, and 0.371, respectively. In VEDAI and DOTA, the mAP(0.5) of LGA-YOLO reaches 0.803 and 0.781, respectively. These metrics not only surpass baseline models but also leading-edge algorithms in the field.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001425414600003  
Indexed Date:  
2025-03-03  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut NUAA, Coll Aerosp Engn, Nanjing 211106, Peoples R China  
Institution Address:  
Beijing Inst Elect Engn, Beijing 100080, Peoples R China  
Institution:  
Institution:  
Beijing Inst Elect Engn

### 788. Lightweight Fire Detection Algorithm Based on Multi-Scale Feature Fusion

摘要: Aiming at the lack of detection accuracy and speed bottleneck of traditional fire detection algorithms, especially for the recognition of small-scale incipient fires and large-scale rapidly spreading fires, this paper proposes a lightweight fire detection algorithm based on multi-scale feature fusion. Firstly, the EDBAN module is designed to replace the C2f module in YOLOv8, to improve the generalization ability and adaptability of the model, especially the accuracy when dealing with multi-scale fire scenarios. Secondly, the original BiFPN structure is improved to adapt to the YOLOv8 model structure, and the Weighted Blend module is designed to weight and fuse the features at each layer, which enhances the feature characterization ability and reduces the risk of missed detection. Finally, the LOTT detection module is further proposed to replace the traditional YOLOv8 detection, which achieves the accuracy and stability of the detection performance while maintaining the lightweight through a series of group convolution and scale adjustment operations. Through experiments on scene-rich fire datasets, the results show that the improved YOLOv8 algorithm reduces the parameter amount by 58.3% and the computation amount by 34.5% based on the baseline model, while the mAP improves by 2.6?percentage points, which basically meets the demand of real-time fire detection.  
摘要:  
针对传统火灾检测算法存在的检测精度不足及速度瓶颈,特别是对于小规模初发火情与大规模迅速蔓延火灾的识别难题,研究提出一种基于多尺度特征融合的轻量级火灾检测算法,设计了EDBAN模块以替代YOLOv8中的C2f模块,提升模型的泛化能力和适应性,尤其是在处理多尺度火灾场景时的精准度。改进原有的BiFPN结构适配YOLOv8模型结构,并设计Weighted Blend模块对各层特征进行加权融合,增强特征的表征能力,降低漏检风险。进一步提出LOTT检测模块,以替代传统的YOLOv8检测,通过一系列组卷积和尺度调整操作,实现了在轻量化的同时保持了检测性能的准确性和稳定性。通过在场景丰富的火灾数据集上进行实验,结果表明,改进的YOLOv8算法在基准模型的基础上参数量减少了58.3%、计算量减少了34.5%,同时mAP提升了2.6个百分点,基本满足火灾实时检测的需求。  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7860928  
Indexed Date:  
2025-02-23  
Institution Address:  
南京审计大学计算机学院, 南京, 江苏 211815, 中国  
Institution Address:  
School of Computer,Nanjing Audit University, Nanjing, Jiangsu 211815, China  
Institution:  
南京审计大学计算机学院  
Institution:  
School of Computer,Nanjing Audit University

### 789. Multi-scale nonlinear edge-based three-phase model for unsupervised

摘要: Unsupervised feature extraction techniques of hyperspectral images (HSIs) have recently drawn significant attention for their excellent performance and efficiency in classification. In some existing methods, the denoising process that reduces the influence of inherent noise is ignored, and the nonlinear edge characteristics and multi-scale features that help to classify still need to be fully considered. To solve these issues, we employ a multi-scale nonlinear edge-based unsupervised three-phase model (UTPM) for hyperspectral feature extraction. Specifically, in the initial phase, a noise-adjusted principal components technique is adopted to lower the noise to improve the performance of the proposed model. Then, a neighbor band grouping technique is designed to reduce redundancy and computational cost with information entropy. Because the information entropy can concretely reflect the importance of different bands in the same group, the inner structure can be maximally preserved. Finally, we utilize a multi-scale feature fusion on kernel low-rank entropic analysis to extract nonlinear edge features and combine it with a convolution algorithm to fuse the elements of multiple scales to improve the classification performance. Compared with several other classical or progressive unsupervised hyperspectral feature extraction algorithms, the classification results on three public HSI datasets validate the effectiveness of UTPM. (c) 2023 Society of Photo-Optical Instrumentation Engineers (SPIE)  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1931-3195  
入藏号:  
WOS:001077860300028  
Indexed Date:  
2023-10-28  
Institution Address:  
Nanjing Univ Posts & Telecommun, Sch Sci, Nanjing, Peoples R China  
Institution Address:  
Beijing Inst Appl Meteorol, Beijing, Peoples R China  
Institution Address:  
Southeast Univ, Sch Math, Nanjing, Peoples R China  
Institution Address:  
Southeast Univ, Jiangsu Prov Key Lab Networked Collect Intelligenc, Nanjing, Peoples R China  
Institution:  
Institution:  
Beijing Inst Appl Meteorol  
Institution:  
Institution:

### 790. Surface defect detection of vehicle light guide plates based on an

摘要: Aiming at the vehicle navigation light guide plate (LGP) image characteristics, such as complex and gradient textures, uneven brightness, and small defects, this paper proposes a visual inspection method for LGP defects based on an improved RetinaNet. First, we use ResNeXt50 with higher accuracy under the same parameters as the backbone network, and propose the lightweight module Ghost\_module to replace the 1 x 1 convolution in the lower half of the ResNeXt\_block. This can reduce the resource parameters and consumption, and speed up training and inference. Second, we propose and use an improved feature pyramid network module to improve the feature fusion network in RetinaNet. It can more effectively fuse the shallow semantic information and high-level semantic information in the backbone feature extraction network, and further improve the detection ability of small target defects. Finally, the defect detection dataset constructed based on the vehicle LGP images collected at a industrial site, and experiments are performed on the vehicle LGP dataset and Aluminum Profile Defect Identification dataset (Aluminum Profile DID). The experimental results show that the proposed method is both efficient and effective. It achieves a better average detection rate of 98.6% on the vehicle LGP dataset. The accuracy and real-time performance can meet the requirements of industrial detection.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
12  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
12  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:000739649000001  
Indexed Date:  
2022-01-12  
Institution Address:  
Zhejiang Sci Tech Univ, Fac Mech Engn & Automat, Inst Automat, Hangzhou 310018, Peoples R China  
Institution:

### 791. SSD Small Target Detection Algorithm Combining Feature Enhancement and

摘要: SSD is a multi-scale target detection algorithm. Due to the lack of semantic information in shallow feature images, the detection accuracy of small targets is low. To solve this problem, a SSD small target detection algorithm, FA-SSD, which combines feature enhancement and self-attention, is proposed. The algorithm constructs a recursive reverse path from deep to shallow based on SSD, which consists of three modules:the deep feature enhancement module uses the contextual information generated from the deep multi-scale feature map and the semantic information of the deepest feature map to enhance the expression ability of the deep feature information;the up-sampling feature enhancement module enhances the semantic information of the up-sampling feature map in the reverse path by enlarging the receptive field of the feature map. The adaptive feature fusion module adaptively fuses adjacent shallow feature images and up-sampling feature images with self-attention mechanism to generate new feature images with strong semantic and precise location information. Experimental results show that on PASCAL VOC and TT100K datasets, the mAP of FA-SSD is up to 92.5% and 80.2%, indicating that this algorithm can enhance the semantic information of shallow feature images and has a good detection effect on small targets in complex scenes.  
摘要:  
SSD是一种多尺度目标检测算法,由于浅层特征图缺乏语义信息,导致小目标的检测准确率低。针对这个问题,提出一种融合特征增强和自注意力的SSD小目标检测算法FA-SSD。该算法在SSD基础上构建一条自深向浅的递归反向路径,此路径包含三个模块:深层特征增强模块利用路径深层多尺度特征图生成的上下文信息和最深层特征图的语义信息,增强深层特征信息的表达能力;上采样特征增强模块通过扩大特征图的感受野,增强反向路径中上采样特征图的语义信息;自适应特征融合模块引入自注意力机制自适应地融合相邻的浅层特征图和上采样特征图,生成新的具有强语义和精确位置信息的特征图。实验结果显示,在PASCAL VOC和TT100K数据集上,FA-SSD的mAP最高达到了92.5%和80.2%,表明该检测算法能够增强浅层特征图的语义信息,对于复杂场景下的小目标有着较好的检测效果。  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
5  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
ISSN:  
1002-8331  
入藏号:  
CSCD:7181655  
Indexed Date:  
2022-06-10  
Institution Address:  
太原理工大学信息与计算机学院, 晋中, 山西 030600, 中国  
Institution Address:  
School of Information and Computer Science, Taiyuan University of Technology, Jinzhong, Shanxi 030600, China  
Institution:  
太原理工大学信息与计算机学院  
Institution:  
School of Information and Computer Science, Taiyuan University of Technology

### 792. Contrastive learning with feature fusion for unpaired thermal infrared

摘要: Colorizing unpaired thermal infrared images is a challenging task that existing methods struggle to perform effectively, often resulting in blurry details and unclear textures. To address these issues, we propose a novel unpaired infrared image colorization method based on contrastive learning with feature fusion. Our proposed method designs the multi-scale feature extraction module, the expansion enhancement module, and the region-aware attention module in the generator. The multi-scale feature extraction module enlarges the model receptive field, achieving more comprehensive feature extraction and fusing feature information of varying scales. The expansion enhancement module enhances the network recognition ability, capturing more semantic contextual information and image details. The region-aware attention module enables the network to obtain more global and local information, adaptively focusing on the feature of image semantic regions. Furthermore, we introduce a new composite loss function that combines adversarial loss, contrastive loss, and perceptual loss, utilizing texture information extracted from the color image to further enhance the details in the generated colorized image and improve image perception. Extensive experimental results on the KAIST dataset and the FLIR dataset show that our proposed method maximally retains tiny optical details in infrared images, such as surface textures, edge contours, and small targets, making the generated color images physically highly consistent with the Ground Truth, with rich and realistic details.  
Web of Science 核心合集中的 "被引频次":  
5  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
0143-8166  
eISSN:  
1873-0302  
入藏号:  
WOS:001087787100001  
Indexed Date:  
2023-11-05  
Institution Address:  
Changchun Univ Sci & Technol, Natl Demonstrat Ctr Expt Elect, Changchun 130022, Jilin, Peoples R China  
Institution:

### 793. Informer Short-Term PV Power Prediction Based on Sparrow Search

摘要: The output power of PV systems is influenced by various factors, resulting in strong volatility and randomness, which makes it difficult to forecast. Therefore, this paper proposes an Informer prediction model based on optimised VMD for predicting short-term PV power. Firstly, the temporal coding of the Informer model is improved and, secondly, the original sequence is decomposed into multiple modal components using VMD, and then optimisation of the results of VMD in conjunction with the optimisation strategy of SSA improves the characteristics of the time series data. Finally, the refined data are fed into the Informer framework for modelling and prediction, utilising the self-attention mechanism and multiscale feature fusion of Informer to precisely forecast PV power. The power of PV prediction data from the SSA-VMD-Informer model and four other commonly used models is compared. Experimental results indicate that the SSA-VMD-Informer model performs exceptionally well in short-term PV power prediction, achieving higher accuracy than traditional methods. As an example, the results of predicting the PV power on 24 April in a region of Xinjiang are 1.3882 for RMSE, 0.8310 for MSE, 1.14 for SDE, and 0.9944 for R2.  
Web of Science 核心合集中的 "被引频次":  
3  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
1996-1073  
入藏号:  
WOS:001255940700001  
Indexed Date:  
2024-07-06  
Institution Address:  
Yunnan Minzu Univ, Sch Elect & Informat Technol, Kunming 650504, Peoples R China  
Institution Address:  
Yunnan Key Lab Unmanned Autonomous Syst, Kunming 650504, Peoples R China  
Institution Address:  
Yunnan Minzu Univ, Lancang Mekong Int Vocat Inst, Kunming 650504, Peoples R China  
Institution:  
Institution:  
Yunnan Key Lab Unmanned Autonomous Syst  
Institution:

### 794. Deep learning assisted vision inspection of resistance spot welds

摘要: Spot welds are extensively used on autobodies and also a key factor affecting the performance of automobiles. Automatic detection of spot welding based on machine vision provides an effective way for car body welding quality control. Considering the traditional image processing methods are greatly disturbed by the environment and have unsatisfying robustness, a network model for small object detection is proposed to detect the position and quality of the spot welding of the car body. Based on the existing You Only Look Once (YOLOv3) model, the proposed model has three novel improvements. Firstly, the lightweight network MobileNetV3 is introduced to replace the backbone network of YOLOv3 to ensure accuracy and real-time performance. Secondly, to improve the model's ability for small object detection, a new feature pyramid network (FPN) with efficient cross-scale connections is proposed, which allows easy and fast multiscale feature fusion. Finally, considering the shortcomings of intersection and union ratio (IoU) loss, complete IoU (CIoU) loss is used to improve convergence speed and regression accuracy. Moreover, novel data augmentation is used to enrich the dataset during the model training. Quantitative results on the spot welding dataset show that the proposed approach achieves successful results for resistance spot welding vision inspection.  
Web of Science 核心合集中的 "被引频次":  
61  
在中国科学引文数据库中的被引频次:  
4  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
65  
ISSN:  
1526-6125  
eISSN:  
2212-4616  
入藏号:  
WOS:000616465900001  
Indexed Date:  
2021-03-04  
Institution Address:  
Shanghai Jiao Tong Univ, State Key Lab Mech Syst & Vibrat, Shanghai 200240, Peoples R China  
Institution:

### 795. Multiscale feature fusion approach to early fault diagnosis in EV power

摘要: The integrity of power batteries is crucial for the safety and reliability of electric vehicles (EVs). In the early stage, the voltage of a failed battery in the thermal runaway vehicle may not be significantly abnormal, which makes it difficult for fault diagnosis methods based on external characteristics to detect potential abnormalities. In this study, an early fault diagnosis method fusing multiscale features is proposed. First, a feature extraction method for internal battery characteristics from actual vehicle data is designed, and the method reveals the complex interactions of internal features, including battery reaction mechanisms, transient behaviors, and longterm voltage trends. On this basis, the inconsistency of the battery pack is measured by the coefficients of variation of the incremental capacity curve features, the approximate internal resistance, and the slope of the charging voltage curve to determine whether a fault has occurred. Subsequently, a multi-classifier fusion fault localization method is proposed, to classify the risk level of battery failure. Comprehensive case studies show that the method can detect several days in advance and accurately locate the first TR batteries, offering superior diagnostic effectiveness and practical applicability in power battery fault diagnosis.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
2352-152X  
eISSN:  
2352-1538  
入藏号:  
WOS:001306292400001  
Indexed Date:  
2024-09-12  
Institution Address:  
South China Univ Technol, Sch Mech & Automot Engn, Guangzhou 510640, Peoples R China  
Institution Address:  
Guangdong Prov Key Lab Automot Engn, Guangzhou 510640, Peoples R China  
Institution:  
Institution:  
Guangdong Prov Key Lab Automot Engn

### 796. Indoor Safety Helmet-Wearing Detection Algorithm Based on Improved

摘要: An improved Yolov4 algorithm was proposed to autodetect indoor safety helmet-wearing. First,a data set dedicated to the indoor safety helmet-wearing detection was self-built for testing and evaluating the algorithm due to a lack of safety helmet-wearing detection experimental data in indoor scenarios. Then,an adaptive recalibration multiscale feature fusion module(ARMFFM) was designed and embedded into the original Yolov4 network to improve the detection accuracy of fuzzy and tiny targets far away from the surveillance camera. In ARMFFM,the features were fused top-down and bottom-up at different scales through depthwise over-parameterized convolutional layers for the fuzzy and tiny objects to obtain the more obvious texture and feature at first. Afterwards,the feature recalibration module strengthened or suppressed each pixel in the fused feature map to make the model precisely detect it to avoid a conflict among the feature maps at different scales. Furthermore,a decoupled detection head replaced the detection head of the original Yolov4 for the individual performances of the location and classification tasks of the indoor safety helmet-wearing detection. Additionally,a Soft-CIoU-NMS post-process algorithm was developed for detecting overlapping targets. The experimental results demonstrated that the accuracy of the improved Yolov4 algorithm in the detection of safety helmet-wearing in indoor scenarios reached 95.1%,about 4.7% higher than that of the original Yolov4. Besides,the detection precision of fuzzy,tiny and overlapping targets was significantly enhanced,proving the superiority of the algorithm for indoor safety helmet-wearing detection.  
摘要:  
为实现智能检测室内作业人员是否佩戴安全帽,提出了一种改进的Yolov4算法.首先,针对目前室内安全帽佩戴状态检测实验数据较为匮乏的问题,自建了一个用于室内场景的安全帽佩戴状态检测数据集.随后,为提升室内监控图像中模糊、微小目标的安全帽佩戴状态检测准确率,设计了自校准多尺度特征融合模块并将其嵌入原Yolov4网络中.该模块首先通过深度超参数化卷积从上至下、从下至上融合不同尺度下的特征,加强待检测目标的特征纹理,使得模型能够检测出这两类目标.再通过特征自校准模块对融合后的特征进行过滤,加强或抑制特征图上的每一像素点,使得模型可以在融合后的特征图上进行精确的检测.此外为加速模型收敛,使用解耦合的检测头替换原Yolov4中的耦合检测头,使目标定位任务与安全帽佩戴状态的分类任务相互独立.最后为提升模型对于重叠目标的检测能力,提出了软性非极大值抑制后处理算法Soft-CIoU-NMS.实验结果表明,该改进的Yolov4模型能够准确地识别出室内作业人员是否佩戴安全帽,准确率达到了95.1%.相比于原Yolov4模型,该模型对位于监控摄像头远端的模糊、微小目标和监控图像中重叠目标的检测能力有明显提升,检测准确率提升了约4.7%,较好地满足了室内场景下作业人员安全帽佩戴状态智能检测的要求.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0493-2137  
入藏号:  
CSCD:7408572  
Indexed Date:  
2023-06-10  
Institution Address:  
北京工业大学信息学部, 北京 100124, 中国  
Institution Address:  
北京交通大学电子信息工程学院, 北京 100044, 中国  
Institution Address:  
Faculty of Information Technology,Beijing University of Technology, Beijing 100124, China  
Institution Address:  
School of Electronic and Information Engineering,Beijing Jiaotong University, Beijing 100044, China  
Institution:  
北京工业大学信息学部  
Institution:  
北京交通大学电子信息工程学院  
Institution:  
Faculty of Information Technology,Beijing University of Technology  
Institution:  
School of Electronic and Information Engineering,Beijing Jiaotong University

### 797. Multitask learning for automatic detection of meniscal injury on 3D knee

摘要: Magnetic resonance imaging (MRI) of the knee is the recommended diagnostic method before invasive arthroscopy surgery. Nevertheless, interpreting knee MRI scans is a time-consuming process that is vulnerable to inaccuracies and inconsistencies. We proposed a multitask learning network MCSNetatt which efficiently introduces segmentation prior features and enhances classification results through multiscale feature fusion and spatial attention modules. The MRI studies and subsequent arthroscopic diagnosis of 259 knees were collected retrospectively. Models were trained based on multitask loss with coronal and sagittal sequences and fused using logistic regression (LR). We visualized the network's interpretability by the gradient-weighted class activation mapping method. The LR model achieved higher area under the curve and mean average precision of medial and lateral menisci than models trained on a single sagittal or coronal sequence. Our multitask model MCSNetat outperformed the single-task model CNet and two clinicians in classification, with accuracy, precision, recall, F1-score of 0.980, 1.000, 0.952, 0.976 for medial and 0.920, 0.905, 0.905, 0.905 for the lateral, respectively. With the assistance of model results and visualized saliency maps, both clinicians showed improvement in their diagnostic performance. Compared to the baseline segmentation model, our model improved dice similarity coefficient and the 95% Hausdorff distance (HD95) of the lateral meniscus for 2.3% and 0.860 mm in coronal images and 4.4% and 2.253 mm in sagittal images. Our multitask learning network quickly generated accurate clinicopathological classification and segmentation of knee MRI, demonstrating its potential to assist doctors in a clinical setting.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0736-0266  
eISSN:  
1554-527X  
入藏号:  
WOS:001367611400001  
PubMed ID:  
39620311  
Indexed Date:  
2024-12-07  
Institution Address:  
Shanghai Jiao Tong Univ, Sch Biomed Engn & Medx Res Inst, 1954 Huashan Rd, Shanghai 200030, Peoples R China  
Institution Address:  
Minist Educ, Engn Res Ctr Digital Med & Clin Translat, Shanghai, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Shanghai Peoples Hosp 6, Sch Med, Dept Orthoped, 600 Yishan Rd, Shanghai 201306, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Shanghai Key Lab Orthopaed Implants & Clin Transla, Shanghai Peoples Hosp 9,Sch Med, R&D Ctr 3D Printing Technol,Dept Orthopaed Surg, Shanghai, Peoples R China  
Institution:  
Institution:  
Minist Educ  
Institution:  
Institution:

### 798. FAPN: Feature Augmented Pyramid Network for polyp segmentation

摘要: Accurate polyp segmentation during colonoscopy examinations can help the clinicians accurately locate polyp areas for further diagnosis or surgeries and thereby decrease the chances of polyps growing into cancer. Although existing approaches can achieve significant improvement by multi-scale feature learning, attention/contextual augmentation, and deep supervision, polyp segmentation is still far from being solved. Actually, enhancing the capability of feature representation may be an excellent way to improve the polyp segmentation performance. From this perspective, we propose a simple but strong framework over feature pyramid network (FPN), called Feature Augmented Pyramid Networks (FAPN), for accurate polyp segmentation with augmented feature representation. Specifically, FAPN consists of three components: Cross-Embedding Module (CEM), Predictive Calibration Module (PCM), and Hierarchical Feature Fusion Module (HFFM). CEM is a two-stage fusion approach that first performs an interactive embedding of multi-level features followed by a second fusion, thus enhancing the fused feature representation. After fusing, PCM leverages the predicted probability maps of each stage (after supervised optimization) to calibrate the fused feature representations, which effectively highlights the regions of interest while avoiding the interference of irrelevant information. Finally, HFFM sequentially combines features from each stage in the top-down pathway, yielding a more robust multi-scale feature representation that allows the framework to segment polyps more accurately. Extensive experiments demonstrate that the proposed network performs favorably against more than a dozen of state-of-the-art methods on five popular polyp segmentation benchmarks.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
17  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
4  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
17  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:000826957800007  
Indexed Date:  
2022-07-23  
Institution Address:  
Univ Elect Sci & Technol China, Sch Informat & Commun Engn, Chengdu 611731, Sichuan, Peoples R China  
Institution Address:  
Chongqing Univ, Sch Microelect & Commun Engn, Chongqing 400044, Peoples R China  
Institution:  
Institution:

### 799. Building extraction based on hyperspectral remote sensing images and

摘要: Hyperspectral remote sensing imaging technology provides assistance in various aspects of daily life through applications such as urban building information statistics and green vegetation estimation. Ensuring the accuracy of automatic thematic information extraction under limited samples is a challenge. In this manuscript, a lightweight semantic segmentation model based on the "encoder-decoder" structure is proposed for extracting buildings from hyperspectral remote sensing images. The proposed model employs the lightweight MobileNet combined with multi scale feature fusion and a group dilated convolution for modelling both shallow and deep spatial and spectral features as the encoder and an efficient combined standardized attention mechanism for selecting the most valuable bands and local information. Extensive experiments reveal that our method produces greater accuracy than state-of-the-art lightweight models in building extraction tasks. We also demonstrated the superiority of our method for insufficient training sample sizes. When only 50% of the samples of the initial training set were used, the mean intersection over union (mIOU) reached 91.90%, 4.5% higher than that of the next best method. For training sets composed of only 16 and 8 images, the mIOU values were 89.42 and 77.11%, respectively, 13.6 and 18 percentage points higher than that of the next best method. According to the visualization of the results, the proposed method obviously outperformed the compared methods. The model proposed in this paper is suitable for accurately extracting buildings from hyperspectral images in situations involving limited training samples.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0045-7906  
eISSN:  
1879-0755  
入藏号:  
WOS:001039836700001  
Indexed Date:  
2023-08-15  
Institution Address:  
Beijing Normal Univ, Adv Inst Nat Sci, Zhuhai 519087, Peoples R China  
Institution Address:  
Univ Sanya, Sch Informat & Intelligence Engn, Sanya 572022, Peoples R China  
Institution Address:  
Univ Sanya, Academician Chen Guoliang Team Innovat Ctr, Sanya 572022, Peoples R China  
Institution:  
Institution:  
Institution:

### 800. SGTL-SUIE: semantic attention-guided transfer learning method for

摘要: The study and analysis of underwater image enhancement are significant in marine engineering and aquatic robotics. The data-driven approach shows good performance in this field. However, the method faces challenges in addressing issues such as low contrast, blurring, and color deviation. Moreover, its performance is constrained by the availability of paired underwater images, making it challenging to capture the nuances among different underwater scenes. To address these challenges, we introduce a semantic attention-guided transfer learning method for stylization underwater image enhancement (SGTL-SUIE). This method enables the generation of multiple stylized and enhanced images from a single distorted underwater image. Within SGTL-SUIE, a style-filtering module is proposed to better bridge the domain gap between distorted images and style reference images. Subsequently, a semantic pairing module further mitigates the domain differences across varying semantics between reference and distorted images, guided by semantic information, producing multiple semantic pairing codes. The transfer enhancement module then takes these semantic pairing codes, co-encoding style features with distorted image features through an encoder. A decoder network subsequently decodes the encoded features into diverse stylized outputs. To validate the proposed method's performance, qualitative and quantitative evaluations were conducted on multiple public datasets. The results demonstrate that the SGTL-SUIE method outperforms many state-of-the-art approaches and enhances the stylistic diversity of generated images. (c) 2024 SPIE and IS&T  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:001377223800054  
Indexed Date:  
2024-12-27  
Institution Address:  
Tianjin Univ Technol, Coll Elect Engn & Automat, Tianjin, Peoples R China  
Institution Address:  
Tianjin Univ Technol, Tianjin Key Lab Complex Syst Control Theory & App, Tianjin, Peoples R China  
Institution Address:  
DEEPINFAR, Tianjin, Peoples R China  
Institution:  
Institution:  
Institution:  
DEEPINFAR

### 801. Prediction of mechanical properties of rolled steel based on

摘要: The mechanical properties of carbon steel sheet directly depend on the chemical composition and process parameters of the steel. There is a complex nonlinear relationship between chemical composition, process parameters, and mechanical properties, and the establishment of a model with the ability to automatically distinguish the importance of different parameters, and to have good prediction accuracy and generalisation in different data sets is one of the important issues in the field of carbon steel mechanical properties prediction. In this paper, we proposed a multi-scale convolutional network model based on dual-attention mechanism (DAM-MSDSC), where the dual-attention module mines from features and channels at the same time, making the model focus on the important features. The multiscale feature fusion module effectively improves the model prediction accuracy by mining features from the output of the middle layer of the network to obtain features containing multiscale fusion features. In hot-rolled steel, for the two parameters of carbon content and temperature, it is confirmed that the change of carbon content has a greater impact on TS, while FTT and CT have a greater impact on YS. Through comparative experiments on different datasets, the proposed model had the best prediction performance, in which the correct rates of YS, TS, and EL are 98 %, 98.7 %, and 99.4 %, respectively, in cold rolled steel, and 97.8 %, 98.6 %, and 99.2 %, respectively, in hot rolled steel.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2352-4928  
入藏号:  
WOS:001321158800001  
Indexed Date:  
2024-10-05  
Institution Address:  
Lanzhou Univ Technol, Sch Comp & Commun, 287 Lan Gong Ping Rd, Lanzhou 730050, Peoples R China  
Institution Address:  
State Key Lab Adv Proc & Recycling Nonferrous Met, Lanzhou 730050, Peoples R China  
Institution Address:  
Lanzhou Univ Technol, Coll Mat Sci & Engn, Lanzhou 730050, Peoples R China  
Institution Address:  
North China Univ Sci & Technol, Sch Nursing & Rehabil, Tangshan 063210, Peoples R China  
Institution:  
Institution:  
State Key Lab Adv Proc & Recycling Nonferrous Met  
Institution:  
Institution:

### 802. Remaining Useful Life Prediction Method Based on Dual-Path Interaction

摘要: In fields such as manufacturing and aerospace, remaining useful life (RUL) prediction estimates the failure time of high-value assets like industrial equipment and aircraft engines by analyzing time series data collected from various sensors, enabling more effective predictive maintenance. However, significant temporal diversity and operational complexity during equipment operation make it difficult for traditional single-scale, single-dimensional feature extraction methods to effectively capture complex temporal dependencies and multi-dimensional feature interactions. To address this issue, we propose a Dual-Path Interaction Network, integrating the Multiscale Temporal-Feature Convolution Fusion Module (MTF-CFM) and the Dynamic Weight Adaptation Module (DWAM). This approach adaptively extracts information across different temporal and feature scales, enabling effective interaction of multi-dimensional information. Using the Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) dataset for comprehensive performance evaluation, our method achieved RMSE values of 0.0969, 0.1316, 0.086, and 0.1148; MAPE values of 9.72%, 14.51%, 8.04%, and 11.27%; and Score results of 59.93, 209.39, 67.56, and 215.35 across four different data categories. Furthermore, the MTF-CFM module demonstrated an average improvement of 7.12%, 10.62%, and 7.21% in RMSE, MAPE, and Score across multiple baseline models. These results validate the effectiveness and potential of the proposed model in improving the accuracy and robustness of RUL prediction.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2076-0825  
入藏号:  
WOS:001340819800001  
Indexed Date:  
2024-10-31  
Institution Address:  
Harbin Engn Univ, Coll Intelligent Syst Sci & Engn, Nantong St, Harbin 150001, Peoples R China  
Institution:

### 803. Graph element detection matching based on Republic of China banknotes

摘要: In view of the fact that there are numerous types of Republic of China banknotes, which often have slight visual differences between different banknote, combined with the issues of mold, burrs or breakage after circulation, the recognition and classification ability of traditional fine-grained image retrieval methods for Republican banknotes is inadequate. To address these issues, this paper proposed a fine-grained retrieval model of Republican banknotes based on multiscale feature fusion. To reduce the time of manual data labeling, YOLOv4 was employed for graph element detection on banknote images, with the main view of banknotes being adopted as the input feature map. EfficientNet-B0 was utilized as the backbone network for retrieval, thereby reducing the burden of redundant information in the network and enhancing network accuracy. In the model, the feature vectors of layers 2, 4, 10, and 15 of the PANet fusion network were utilized to generate a global feature vector library, improving the banknote matching retrieval capability. Furthermore, the feature vectors were clustered using adaptive K-means to simplify the matching time and computation. The experimental results demonstrated that the proposed model achieved an accuracy of 89.6%, improving the retrieval accuracy by 10 percentage points compared to using the original image of banknotes as the input image. The improved model exhibited better classification performance, less inference time cost, and fine classification of banknotes. These results could meet the practical requirements of industry.  
摘要:  
民国纸币种类数量众多,不同纸币类别间的视觉差异小,部分纸币经过流通后发霉、毛边以及破损。针对传统的细粒度图像检索方法对民国纸币识别分类能力差的问题,提出了一种基于多尺度特征融合的民国纸币细粒度检索模型。在使用YOLOv4对纸币图像做图元素检测,减少手动标记数据时间的基础上,利用纸币主景图作为输入特征图,使用EfficientNet-B0作为主干网络进行检索,减少了冗余信息对网络的负担,提升了网络的精度。在模型中,使用PANet融合网络的第2,4,10和15层的特征向量,生成全局特征向量库,提升了纸币匹配检索能力,并使用自适应K均值对特征向量进行聚类,简化了匹配的时间与计算量。实验结果表明,该模型准确率达到了89.6%,相比于使用纸币原图作为输入图像提升了10个百分点,提高了检索精度。改进后的模型分类效果更好,推理时间成本更少,实现了纸币的精细化分类。满足工业实际要求。  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2095-302X  
入藏号:  
CSCD:7499427  
Indexed Date:  
2023-10-14  
Institution Address:  
上海师范大学信息与机电工程学院, 上海 200234, 中国  
Institution Address:  
School of Information and Electromechanical Engineering, Shanghai Normal University, Shanghai 200234, China  
Institution:  
上海师范大学信息与机电工程学院  
Institution:  
School of Information and Electromechanical Engineering, Shanghai Normal University

### 804. Sorting of <i>Mountage Cocoons</i> Based on MobileSAM and Target

摘要: The classification of silkworm cocoons is essential prior to silk reeling and serves as a key step in improving the quality of raw silk. At present, cocoon classification mainly relies on manual sorting, which is labor-intensive and inefficient. In this paper, a cocoon detection algorithm S-YOLOv8\_c based on the cooperation of MobileSAM and YOLOv8 for the mountage cocoons was proposed. The MobileSAM with a designed area thresholding algorithm was used for the semantic segmentation of mountage cocoon images, which could mitigate the effect of complex backgrounds and maximize the discriminability of cocoon features. Subsequently, the BiFPN was added to the neck of YOLOv8 to improve the multiscale feature fusion capability. The loss function was replaced with the WIoU, and a dynamic non-monotonic focusing mechanism was introduced to improve the generalization ability. In addition, the GAM was incorporated into the head to focus on detailed cocoon information. Finally, the S-YOLOv8\_c achieved a good detection accuracy on the test set, with a mAP of 95.8%. Furthermore, to experimentally validate the sorting ability, we deployed the proposed model onto the self-developed Cartesian coordinate automatic cocoon harvester, which indicated that it would effectively meet the requirements of accurate and efficient cocoon sorting.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
2077-0472  
入藏号:  
WOS:001211435400001  
Indexed Date:  
2024-05-10  
Institution Address:  
Shandong Agr Univ, Coll Mech & Elect Engn, Tai An 271018, Peoples R China  
Institution Address:  
Sericulture Technol Promot Stn Guangxi Zhuang Auto, Nanning 530000, Peoples R China  
Institution Address:  
Guangxi Key Lab Silkworm Genet Improvement & Effic, Nanning 530000, Peoples R China  
Institution Address:  
Shandong Engn Res Ctr Intelligent Agr Equipment, Tai An 271018, Peoples R China  
Institution Address:  
Shandong Key Lab Hort Machinery & Equipment, Tai An 271018, Peoples R China  
Institution:  
Institution:  
Sericulture Technol Promot Stn Guangxi Zhuang Auto  
Institution:  
Guangxi Key Lab Silkworm Genet Improvement & Effic  
Institution:  
Shandong Engn Res Ctr Intelligent Agr Equipment  
Institution:  
Shandong Key Lab Hort Machinery & Equipment

### 805. Semantic feature-constrained multitask siamese network for building

摘要: In the field of remote sensing applications, semantic change detection (SCD) simultaneously identifies changed areas and their change types by jointly conducting bitemporal image classification and change detection. It facilitates change reasoning and provides more application value than binary change detection (BCD), which offers only a binary map of the changed/unchanged areas. In this study, we propose a multitask Siamese network, named the semantic feature-constrained change detection (SFCCD) network, for building change detection in bitemporal high-spatial-resolution (HSR) images. SFCCD conducts feature extraction, semantic segmentation and change detection simultaneously, where change detection and semantic segmentation are the main and auxiliary tasks, respectively. For the segmentation task, ResNet50 is used to conduct image feature extraction, and the extracted semantic features are provided to execute the change detection task via a series of jump connections. For the change detection task, a global channel attention (GCA) module and a multiscale feature fusion (MSFF) module are designed, where high-level features offer training guidance to the low-level feature maps, and multiscale features are fused with multiple convolutions that possess different receptive fields. In bitemporal HSR images with different view angles, high-rise buildings have different directional height displacements, which generally cause serious false alarms for common change detection methods. However, known public building change detection datasets often lack buildings with height displacement. We thus create the Nanjing Dataset (NJDS) and design the aforementioned network structures and modules to target this issue. Experiments for method validation and comparison are conducted on the NJDS and two additional public datasets, i.e., the WHU Building Dataset (WBDS) and Google Dataset (GDS). Ablation experiments on the NJDS show that the joint utilization of the GCA and MSFF modules performs better than several classic modules, including atrous spatial pyramid pooling (ASPP), efficient spatial pyramid (ESP), channel attention block (CAB) and global attention upsampling (GAU) modules, in dealing with building height displacement. Furthermore, SFCCD achieves higher accuracy in terms of the OA, recall, F1-score and mIoU measures than several state-of-the-art change detection methods, including deeply supervised image fusion network (DSIFN), the dual-task constrained deep Siamese convolutional network (DTCDSCN), and multitask U-Net (MTU-Net).  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
56  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
56  
ISSN:  
0924-2716  
eISSN:  
1872-8235  
入藏号:  
WOS:000800382100006  
Indexed Date:  
2022-06-08  
Institution Address:  
Nanjing Normal Univ, Key Lab Virtual Geog Environm, Minist Educ, Nanjing 210023, Peoples R China  
Institution Address:  
Nanjing Normal Univ, Sch Geog, Nanjing 210023, Peoples R China  
Institution Address:  
Jiangsu Ctr Collaborat Innovat Geog Informat Resou, Nanjing 210023, Peoples R China  
Institution Address:  
State Key Lab Cultivat Base Geog Environm Evolut J, Nanjing 210023, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100101, Peoples R China  
Institution:  
Institution:  
Institution:  
Jiangsu Ctr Collaborat Innovat Geog Informat Resou  
Institution:  
State Key Lab Cultivat Base Geog Environm Evolut J  
Institution:

### 806. MSDFNet: multi-scale detail feature fusion encoder-decoder network for

摘要: Currently available thermal image depth estimation methods are difficult to efficiently extract fine multi-scale feature information from thermal images and suffer from the problem of blurring details at the edges of the estimated depth map. To address these challenges, this paper proposes MSDFNet, a multi-scale detail feature fusion encoder-decoder network, for self-supervised monocular thermal image depth estimation. The model is based on a channel expansion hourglass residual lightweight feature encoder, which can capture rich and fine-grained multi-scale feature information with low computational effort. MSDFNet utilizes a detail feature weight evaluation decoder to fuse cross-scale features and reevaluate the importance of each feature, thereby emphasizing critical edge information at multiple scales. Additionally, MSDFNet incorporates a depth consistency loss function, which provides self-supervised signals for the detailed features of thermal images and improves the optimization of network performance. The method is applied to the ViViD++ and MS2 datasets and achieves state-of-the-art depth estimation performance compared to existing state-of-the-art algorithms. In the Indoor Dark scenario of the ViViD++ dataset, the Abs Rel, Sq Rel, RMSE, and RMSE log error metric values of MSDFNet are reduced by 6.71%, 11.92%, 9.09%, and 5.73%, respectively, while the accuracy metric values delta < 1.25(i), i = 1,2,3 were improved by 4.18%, 1.13%, and 0.2%, respectively. In addition, MSDFNet proves its excellent generalization ability on the MS2 dataset. The Abs Rel and RMSE error values in the night scene are reduced by 45.6% and 30.09%, respectively, and the accuracy delta < 1.25(i), i = 1,3 is improved by 20.95% and 1.33%, respectively. The Abs Rel and RMSE values in the rainy day scenario are reduced by 1.33% and 1.21%, respectively, and the accuracy delta < 1.25(i),i = 1,3 is improved by 0.24% and 0.83%, respectively.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001370085200001  
Indexed Date:  
2024-12-13  
Institution Address:  
Shanghai Publishing & Printing Coll, Shanghai 200093, Peoples R China  
Institution Address:  
Univ Shanghai Sci & Technol, Coll Publishing, Shanghai 200093, Peoples R China  
Institution:  
Shanghai Publishing & Printing Coll  
Institution:

### 807. Heterogeneous remote sensing image change detection network based on

摘要: With the rapid advancement of remote sensing (RS) earth observation technology, it is possible to monitor land use/land cover change with multi-modal RS data. The challenge of heterogeneous RS image change detection (HRSI-CD) is determining how to unify the feature space of images such that they are comparable. However, the majority of existing methods often extract features from the bi-temporal heterogeneous images (BTHIs) in order to acquire shared features. During the process of feature extraction, the correlation and interactivity between images captured at different times are disregarded, which leads to the information loss during images conversion and greatly limits the performance of CD. The correlation mechanism between heterogeneous images needs to be further explored and utilized. A new multi-scale cross-modal fusion network based on style-transfer (MCFNet) is proposed in this paper to solve these problems. The idea of MCFNet involves the restructuring of the content and style representation of heterogeneous images. The network employs content encoders and style encoders to extract multi-scale content and style features from BTHIs. It also enforces consistency in style representation by applying similarity style constraints at the minimum scale, ensuring that related pixels have consistent styles. On this basis, we propose a cross-modal feature cross- fusion mechanism that integrates the content and style features of the two images, effectively exploiting the association between the heterogeneous images. Furthermore, a multi-scale feature fusion method is proposed to effectively aggregate the content and style features of images across various scales. This technique aims to achieve precise style transfer while maintaining the semantic information of the images. We verify the proposed MCFNet on three HRSI-CD datasets. The experimental results demonstrate that the proposed method possesses significant advantages in comparison to the prevailing methods. On three datasets, MCFNet outperforms other advanced algorithms by at least 4% in terms of the Kappa indicator.  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1568-4946  
eISSN:  
1872-9681  
入藏号:  
WOS:001420161200001  
Indexed Date:  
2025-02-18  
Institution Address:  
Liaoning Normal Univ, Sch Geog Sci, Dalian 116029, Peoples R China  
Institution Address:  
Liaoning Normal Univ, Sch Comp Sci & Artificial Intelligence, Dalian 116029, Peoples R China  
Institution:  
Institution:

### 808. GVC-YOLO: A Lightweight Real-Time Detection Method for Cotton

摘要: Cotton aphids (Aphis gossypii Glover) pose a significant threat to cotton growth, exerting detrimental effects on both yield and quality. Conventional methods for pest and disease surveillance in agricultural settings suffer from a lack of real-time capability. The use of edge computing devices for real-time processing of cotton aphid-damaged leaves captured by field cameras holds significant practical research value for large-scale disease and pest control measures. The mainstream detection models are generally large in size, making it challenging to achieve real-time detection on edge computing devices with limited resources. In response to these challenges, we propose GVC-YOLO, a real-time detection method for cotton aphid-damaged leaves based on edge computing. Building upon YOLOv8n, lightweight GSConv and VoVGSCSP modules are employed to reconstruct the neck and backbone networks, thereby reducing model complexity while enhancing multiscale feature fusion. In the backbone network, we integrate the coordinate attention (CA) mechanism and the SimSPPF network to increase the model's ability to extract features of cotton aphid-damaged leaves, balancing the accuracy loss of the model after becoming lightweight. The experimental results demonstrate that the size of the GVC-YOLO model is only 5.4 MB, a decrease of 14.3% compared with the baseline network, with a reduction of 16.7% in the number of parameters and 17.1% in floating-point operations (FLOPs). The mAP@0.5 and mAP@0.5:0.95 reach 97.9% and 90.3%, respectively. The GVC-YOLO model is optimized and accelerated by TensorRT and then deployed onto the embedded edge computing device Jetson Xavier NX for detecting cotton aphid damage video captured from the camera. Under FP16 quantization, the detection speed reaches 48 frames per second (FPS). In summary, the proposed GVC-YOLO model demonstrates good detection accuracy and speed, and its performance in detecting cotton aphid damage in edge computing scenarios meets practical application needs. This research provides a convenient and effective intelligent method for the large-scale detection and precise control of pests in cotton fields.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
2072-4292  
入藏号:  
WOS:001304707900001  
Indexed Date:  
2024-09-08  
Institution Address:  
Henan Agr Univ, Coll Informat & Management Sci, Zhengzhou 450002, Peoples R China  
Institution Address:  
Xinjiang Acad Agr Sci, Inst Plant Protect, Urumqi 830091, Peoples R China  
Institution Address:  
Chinese Acad Agr Sci, Inst Plant Protect, Beijing 100193, Peoples R China  
Institution Address:  
Natl Agr Technol Serv Ctr, Div Pest Monitoring & Forecasting, Beijing 100125, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Natl Agr Technol Serv Ctr

### 809. Difficult Airway Assessment Based on Multi-View Metric Learning

摘要: The preoperative assessment of difficult airways is of great significance in the practice of anesthesia intubation. In recent years, although a large number of difficult airway recognition algorithms have been investigated, defects such as low recognition accuracy and poor recognition reliability still exist. In this paper, we propose a Dual-Path Multi-View Fusion Network (DMF-Net) based on multi-view metric learning, which aims to predict difficult airways through multi-view facial images of patients. DMF-Net adopts a dual-path structure to extract features by grouping the frontal and lateral images of the patients. Meanwhile, a Multi-Scale Feature Fusion Module and a Hybrid Co-Attention Module are designed to improve the feature representation ability of the model. Consistency loss and complementarity loss are utilized fully for the complementarity and consistency of information between multi-view data. Combined with Focal Loss, information bias is effectively avoided. Experimental validation illustrates the effectiveness of the proposed method, with the accuracy, specificity, sensitivity, and F1 score reaching 77.92%, 75.62%, 82.50%, and 71.35%, respectively. Compared with methods such as clinical bedside screening tests and existing artificial intelligence-based methods, our method is more accurate and reliable and can provide a reliable auxiliary tool for clinical healthcare personnel to effectively improve the accuracy and reliability of preoperative difficult airway assessments. The proposed network can help to identify and assess the risk of difficult airways in patients before surgery and reduce the incidence of postoperative complications.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
eISSN:  
2306-5354  
入藏号:  
WOS:001276715600001  
PubMed ID:  
39061785  
Indexed Date:  
2024-08-01  
Institution Address:  
Southwest Jiaotong Univ, Sch Comp & Artificial Intelligence, Chengdu 611756, Peoples R China  
Institution Address:  
Sichuan Univ, West China Hosp, Gen Practice Med Ctr, Chengdu 610044, Peoples R China  
Institution Address:  
Sichuan Univ, West China Hosp, Dept Anesthesiol, Chengdu 610044, Peoples R China  
Institution:  
Institution:  
Institution:

### 810. Attentive encoder-decoder networks for crowd counting

摘要: Crowd counting that aims to estimate the crowd density has recently made significant progress but remains an unsolved problem due to several challenges. In this paper, we propose an Attentive Encoder-Decoder Network (AEDNet) to overcome the notorious scale-variation problem in crowd counting. Our major contributions can be summarized in three aspects. First, we design an Attentive Feature Refinement (AFR) block in the encoder to adaptively extract multi-scale features. AFR compares the spatial information in different scales through the attention mechanism and then adaptively assign importance weights to each point, which highlights the distinctive roles in multi-scale feature extraction. Second, we develop a Separable Non-local Fusion (SNF) block in the decoder with the self-attention mechanism to aggregate multi-scale features from different layers, which not only achieves the sufficient feature fusion by capturing long-range dependencies, but also vastly reduces the computation cost compared to the original non-local operation. Third, we propose a Regional MSE (R-MSE) loss to tackle the pixel-isolation problems in regular MSE loss. To demonstrate the effectiveness of the proposed AEDNet, we conduct extensive experiments on four widely-used crowd counting datasets, and our AEDNet consistently achieves the state-of-the-art performance. (c) 2021 Elsevier B.V. All rights reserved.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000796199200006  
Indexed Date:  
2022-05-31  
Institution Address:  
Beihang Univ, Sch Elect & Informat Engn, Beijing, Peoples R China  
Institution Address:  
Beihang Univ, Sch Automat Sci & Elect Engn, Beijing, Peoples R China  
Institution Address:  
Incept Inst Artificial Intelligence, Beijing, Peoples R China  
Institution Address:  
Beihang Univ, Sch Astronaut, Beijing, Peoples R China  
Institution:  
Institution:  
Institution:  
Incept Inst Artificial Intelligence  
Institution:

### 811. Sparse Embedded Convolution Based Dual Feature Aggregation 3D Object

摘要: The algorithm design of compatible detection speed and accuracy based on LiDAR point clouds is a challenging issue in various practical applications of 3D object detection, including the field of autonomous driving. This paper designs a single-stage object detection algorithm that is lightweight and compatible with detection speed and accuracy for the above issue. To achieve these objectives, we propose a framework for a 3D object detection algorithm using a single-stage detection network as the backbone network. Firstly, we design a dual feature extraction module to reduce the occurrence of vehicle miss and error detection problems. Then, we use a multi-scale feature fusion scheme to fuse feature information with different scales. Furthermore, we design a data enhancement scheme suitable for this network architecture. Experimental results in the KITTI dataset show that the proposed method achieves improvement ratios of 38.5% for the detection speed and 2.88% similar to\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\sim $$\end{document} 13.65% in terms of the average precision of vehicle detection compared to the existing algorithm based on single-stage object detection (SECOND).  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1370-4621  
eISSN:  
1573-773X  
入藏号:  
WOS:001162617100003  
Indexed Date:  
2024-02-24  
Institution Address:  
Guangxi Normal Univ, Sch Elect & Informat Engn, Guilin 541004, Guangxi, Peoples R China  
Institution Address:  
Guangxi Normal Univ, Educ Dept Guangxi, Key Lab Nonlinear Circuits & Opt Commun, Guilin 541004, Guangxi, Peoples R China  
Institution Address:  
Guangxi Normal Univ, Sch Elect & Informat Engn, Guangxi Key Lab Brain inspired Comp & Intelligent, Guilin 541004, Guangxi, Peoples R China  
Institution Address:  
Guangxi Normal Univ, Coll Elect & Informat Engn, Guilin 541004, Guangxi, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 812. Dual-Modal Feature Fusion Semantic Segmentation of RGB-D

摘要: The existing RGB image semantic segmentation network for complex indoor scenes is susceptible to factors such as color and lighting, while it is also challenging to integrate dual-modal features effectively. Regarding the issue indicated above, this paper proposes an attention mechanism bimodal fusion network(AMBFNet) that adopts an encoder-decoder structure. In the first phase, building the bimodal fusion network structure(AMBF) is carried out to reasonably allocate the location and channel information of the features at each stage of the encoding branch. And then, designing the DA-context module is implemented to merge the context information. Finally, the multi-scale feature maps are cross-layer fused through the decoder to reduce the problem of misrecognition between classes and the loss of small-scale targets in the prediction results. The test results on the two public datasets of SUN RGB-DNYU and Depth v2(NYUDV2) show the consequence that compared with the more advanced RGB-D semantic segmentation network such as the RedNet, ACNet and ESANet, under the same hardware conditions, the network proposed in this paper has better segmentation performance. At the same time, the MIoU reaches 47.9% and 50.0%, respectively.  
摘要:  
针对复杂室内场景中,现有RGB图像语义分割网络易受颜色、光照等因素影响以及RGB-D图像语义分割网络难以有效融合双模态特征等问题,提出一种基于注意力机制的RGB-D双模态特征融合语义分割网络AMBFNet(attention mechanism bimodal fusion network)。该网络采用编-解码器结构,首先搭建双模态特征融合结构(AMBF)来合理分配编码支路各阶段特征的位置与通道信息,然后设计双注意感知的上下文(DA-context)模块以合并上下文信息,最后通过解码器将多尺度特征图进行跨层融合,以减少预测结果中类间误识别和小尺度目标丢失问题。在SUN RGB-DNYU和NYU Depth v2(NYUDV2)两个公开数据集上的测试结果表明,相较于残差编解码(RedNet)、注意力互补网络(ACNet)、高效场景分析网络(ESANet)等目前较先进的RGB-D语义分割网络,在同等硬件条件下,该网络具有更好的分割性能,平均交并比(MIoU)分别达到了47.9%和50.0%。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7450745  
Indexed Date:  
2023-07-28  
Institution Address:  
西南科技大学信息工程学院;;西南科技大学, ;;特殊环境机器人技术四川省重点实验室, 绵阳;;绵阳, 四川;;四川 621010;;621010, 中国  
Institution Address:  
School of Information Engineering,Southwest University of Science and Technology;;Southwest University of Science and Technology, ;;Robot Technology Used for Special Environment Key Laboratory of Sichuan Province, Mianyang;;Mianyang, Sichuan;;Sichuan 621010;;621010  
Institution:  
西南科技大学信息工程学院;;西南科技大学  
Institution:  
School of Information Engineering,Southwest University of Science and Technology;;Southwest University of Science and Technology

### 813. EAFF-Net: Efficient attention feature fusion network for dual-modality

摘要: The pedestrian detection network utilizing a combination of infrared and visible image pairs can improve detection accuracy by fusing their complementary information, especially in challenging illumination conditions. However, most existing dual-modality methods only focus on the effectiveness of feature maps between different modalities while neglecting the issue of redundant information in the modalities. This oversight often affects the detection performance in low illumination conditions. This paper proposes an efficient attention feature fusion network (EAFF-Net), which suppresses redundant information and enhances the fusion of features from dualmodality images. Firstly, we design a dual-backbone network based on CSPDarknet53 and combine with an efficient partial spatial pyramid pooling module (EPSPPM), improving the efficiency of feature extraction in different modalities. Secondly, a feature attention fusion module (FAFM) is built to adaptively weaken modal redundant information to improve the fusion effect of features. Finally, a deep attention pyramid module (DAPM) is proposed to cascade multi-scale feature information and obtain more detailed features of small targets. The effectiveness of EAFF-Net in pedestrian detection has been demonstrated through experiments conducted on two public datasets.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1350-4495  
eISSN:  
1879-0275  
入藏号:  
WOS:001409241100001  
Indexed Date:  
2025-02-07  
Institution Address:  
Fuzhou Univ, Sch Mech Engn & Automat, Xueyuan Rd 2, Fuzhou 350000, Fujian, Peoples R China  
Institution:

### 814. Automatic detection and recognition of electric vehicle helmet based on

摘要: Aiming at the problems of low detection precision,poor robustness,and imperfect related systems in the current small object detection of electric vehicle helmet,an electric vehicle helmet detection model was proposed based on improved YOLOv5s algorithm.In the proposed model,Convolutional Block Attention Module (CBAM) and Coordinate Attention (CA) module were introduced,and the improved Non-Maximum Suppression (NMS)-Distance Intersection over Union-Non Maximum Suppression (DIoU-NMS) was used.At the same time,multi-scale feature fusion detection was added and densely connected network was combined to improve feature extraction effect.Finally,a helmet detection system for electric vehicle drivers was established.The improved YOLOv5s algorithm had the mean Average Precision (mAP) increased by 7.1 percentage points when the Intersection over Union (IoU) is 0.5,and Recall increased by 1.6 percentage points compared with the original YOLOv5s on the self-built electric vehicle helmet wearing dataset.Experimental results show that the improved YOLOv5s algorithm can better meet the requirements for detection precision of electric vehicles and the helmets of their drivers in actual situations,and reduce the incidence rate of electric vehicle traffic accidents to a certain extent.  
摘要:  
针对目前电动车头盔小目标检测的精度低、鲁棒性差,相关系统不完善等问题,提出了基于改进YOLOv5s的电动车头盔检测算法。所提算法引入卷积块注意力模块(CBAM)和协调注意力(CA)模块,采用改进的非极大值抑制(NMS),即DIoU-NMS(Distance Intersection over Union-Non Maximum Suppression);同时增加多尺度特征融合检测,并结合密集连接网络改善特征提取效果;最后,建立了电动车驾驶人头盔检测系统。在自建的电动车头盔佩戴数据集上,当交并比(IoU)为0.5时,所提算法的平均精度均值(mAP)比原始YOLOv5s提升了7.1个百分点,召回率(Recall)提升了1.6个百分点。实验结果表明,所提改进的YOLOv5s算法更能满足在实际情况中对电动车及驾驶员头盔的检测精度要求,一定程度上降低了电动车交通事故的发生率。  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
2  
ISSN:  
1001-9081  
入藏号:  
CSCD:7457722  
Indexed Date:  
2023-08-04  
Institution Address:  
西安科技大学通信与信息工程学院, 西安, 陕西 710600, 中国  
Institution Address:  
College of Communication and Information Technology,Xi'an University of Science and Technology, Xi'an, Shaanxi 710600, China  
Institution:  
西安科技大学通信与信息工程学院  
Institution:  
College of Communication and Information Technology,Xi'an University of Science and Technology

### 815. Multiscale Semantic Fusion-Guided Fractal Convolutional Object Detection

摘要: Optical remote sensing object detection is a challenging task, because of the complex background interference, ambiguous appearances of tiny objects, densely arranged circumstances, and multiclass object with vaster scale variances and irregular aspect ratios. The performance of object detection is seriously restricted. Thus, in this article, inspired by the anchor-free object detection framework, and aiming to solve these difficulties to improve the optical remote sensing object detection performance, a powerful one-stage detector of multiscale semantic fusion-guided fractal convolution network (MSFC-Net) is proposed. First, facing these strong-coupled semantic relations in each complex scene, a compound semantic feature fusion (CSFF) way is designed for generating an effective semantic description, which is a benefit to pixel-wise object center point interpretation. In addition, it can be easily extended into a semantic segmentation task. Second, in view of accurate multiclass pixel-wise center point predictions based on an effective compound semantic description, a novel fractal convolution (FC) regression layer is designed, which adaptively achieves the regression of multiscale bounding boxes (bboxes) with irregular aspect ratio under no priori information. Third, related to the set up FC regression layer, a specific hybrid loss is designed to make the proposed MSFC-Net converge better. Finally, the extensive experiments on challenge data sets of large-scale dataset for object detection in aerial images (DOTA) and object detection in optical remote sensing images (DIOR) datasets are carried out, and comparisons indicate that the proposed MSFC-Net can perform the remarkable performance than other state-of-the-art one-stage detectors, as it can reach 80.26% mean average precision (mAP) and 79.33% mF1 on DOTA and 70.08% mAP and 73.45% mF1 on DIOR. Then, our work is available at https://github.com/ZhAnGToNG1/MSFC-Net.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
50  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
51  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000732807900001  
Indexed Date:  
2021-12-28  
Institution Address:  
Beijing Inst Technol, Beijing Key Lab Embedded Real Time Informat Proc, Beijing 100081, Peoples R China  
Institution Address:  
Commun Univ China, Engn Ctr Digital Audio & Video, Beijing 100024, Peoples R China  
Institution Address:  
Peking Univ, Sch Elect Engn & Comp Sci, Beijing 100871, Peoples R China  
Institution:  
Institution:  
Institution:

### 816. Remainder particles detection of spacecraft based on

摘要: Remainder particles in closed electronic equipment equipped in spacecraft bring huge hidden danger to the flight safety of spacecraft.Since remainder particles are in small size,and even the morphological structure of the remainder particles is highly similar to the general components in equipment,and remainder particles are easily covered by other components,the current methods used to detect remainder particles can cause false detection and missed detection frequently.To resolve these problems,a Remainder Particle Detection Network (RPDN) was proposed to detect remainder particles in closed electronic equipment based on convolution-inverted residual and combined attention mechanism.A convolution-inverted residual module was built to ensure the integrity of the remainder particles' fine-grained feature.Then,the combined attention mechanism was proposed to enhance the representativeness of remainder particles feature.The objects were predicted from multiple dimensions by combining multi-scale feature fusion module and object detection layer.The experimental results showed that RPDN had achieved good effect in all evaluation indicators,the mAP of the proposed method reached to 92.16%,and the detection efficiency reached 13FPS.It realized efficient and accurate detection of remainder particles in closed electronic equipment equipped in spacecraft.  
摘要:  
航天器密闭电子设备内腔多余物给航天器飞行安全带来了巨大隐患。由于多余物体积小、与设备内常规组件形态结构相似且易被其他组件遮挡,采用现有的方法对其进行检测时误检、漏检频发。为解决上述问题,提出一种基于卷积一反残差和组合注意力机制的航天器密闭电子设备多余物检测网络RPDN。首先,网络通过构建卷积一反残差模块,保证了多余物细粒度特征的完整性;其次,设计组合注意力机制,增强了多余物特征的表征能力;最后,结合多尺度特征融合模块与目标检测层从多维度进行目标预测。实验结果表明RPDN在各项评价指标上均取得了良好的效果,mAP达到92.16%,检测效率达到了13FPS,实现了航天器密闭电子设备内腔多余物高效、精准检测。  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-5911  
入藏号:  
CSCD:7652679  
Indexed Date:  
2024-04-26  
Institution Address:  
南京航空航天大学计算机科学与技术学院, 南京, 江苏 210001, 中国  
Institution Address:  
南京航空航天大学机电学院, 南京, 江苏 210001, 中国  
Institution Address:  
南京航空航天大学计算机科学与技术学院;;南京航空航天大学机电学院, ;;, 南京;;南京, 江苏;;江苏 210001;;210001, 中国  
Institution Address:  
College of Computer Science and Technology,Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 210001, China  
Institution Address:  
College of Mechanical and Electrical Engineering,Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 210001, China  
Institution Address:  
College of Computer Science and Technology,Nanjing University of Aeronautics and Astronautics;;College of Mechanical and Electrical Engineering,Nanjing University of Aeronautics and Astronautics, ;;, Nanjing;;Nanjing, ;; 210001;;210001  
Institution:  
南京航空航天大学计算机科学与技术学院  
Institution:  
南京航空航天大学机电学院  
Institution:  
南京航空航天大学计算机科学与技术学院;;南京航空航天大学机电学院  
Institution:  
College of Computer Science and Technology,Nanjing University of Aeronautics and Astronautics  
Institution:  
College of Mechanical and Electrical Engineering,Nanjing University of Aeronautics and Astronautics  
Institution:  
College of Computer Science and Technology,Nanjing University of Aeronautics and Astronautics;;College of Mechanical and Electrical Engineering,Nanjing University of Aeronautics and Astronautics

### 817. Real-time vehicle detection algorithm based on a lightweight

摘要: A vehicle detection algorithm is of great significance for automatic driving technology. Current vehicle detection algorithms suffer from the complex structure, high configuration of hardware requirements, and the difficulty to apply to mobile terminal equipment. In order to solve these issues, this paper proposes an improved YOLOv5 algorithm, named YOLOv5n-L, for lightweight. First, a depthwise separable convolution and a C3Ghost module are used to replace several C3 modules to reduce the model parameters and improve the detection speed. Then a Squeeze-and-Excitation attention mechanism is integrated into backbone network to improve the accuracy of the algorithm and suppress the environmental interference. Finally, a bidirectional feature pyramid network is used for multi-scale feature fusion to enrich feature information and improve the feature extraction ability of the proposed algorithm. The experimental results demonstrate that compared with the original algorithm, the model weight is reduced by 40 % to only 2.3 M. The mean average precision (mAP@0.5) is increased by 1.7 %. The detection speed reaches 80 FPS, which could accurately detect vehicle targets in real-time.  
Web of Science 核心合集中的 "被引频次":  
45  
在 BIOSIS Citation Index 中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
被引频次合计:  
47  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:000884391200005  
Indexed Date:  
2022-11-25  
Institution Address:  
Changchun Univ Sci & Technol, Dept Elect & Informat Engn, Jilin 130022, Jilin, Peoples R China  
Institution Address:  
Chinese Acad Sci, Changchun Inst Opt Fine Mech & Phys, Changchun 130033, Peoples R China  
Institution Address:  
Univ Cambridge, Dept Engn, Photon & Sensors Grp, Cambridge CB3 0FA, England  
Institution Address:  
Beihang Univ, Sch Instrumentat & Optoelect Engn, Beijing 100191, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 818. Improved YOLOv7 algorithm for flame detection in complex urban

摘要: To address the problems of flame detection, such as difficulties in detecting flames and poor performance in complex urban environments, an improved YOLOv7-based flame detection algorithm for complex urban scenarios is proposed. The proposed algorithm increases multi-scale feature fusion and introduces a 160 x 160 detection scale, which improves the detection capability of small target flames. Additionally, the 3 x 3 convolutions in the backbone feature extraction module of YOLOv7 are replaced with deformable convolutions (Deformable Convolution Networks v2, DCNv2), which better accommodate varying input feature map shapes and enhance the network's learning ability in complex scenarios. Furthermore, the Convolutional Block Attention Module (CBAM) is embedded in the feature extraction module to strengthen the response to relevant features, further improving the algorithm's performance in dynamic environments. The K-means++ algorithm is used to re-cluster the anchor boxes, enhancing the algorithm's ability to predict target sizes and locations. The modified YOLOv7 algorithm achieves a mean Average Precision (mAP@0.5) of 97.1%, an improvement of 4.9 percentage points. Experimental results demonstrate that the improved YOLOv7 algorithm significantly enhances flame detection in complex urban scenarios.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2631-8695  
入藏号:  
WOS:001446581700001  
Indexed Date:  
2025-03-23  
Institution Address:  
Beijing Univ Civil Engn & Architecture, Sch Mech Elect & Vehicle Engn, Beijing 100044, Peoples R China  
Institution:

### 819. Efficient three-dimensional point cloud object detection based on

摘要: Lidar-based 3D object detection and classification is a critical task for autonomous driving. However, inferencing from exceedingly sparse 3D data in real-time is a formidable challenge. Complex-YOLO solves the problem of point cloud disorder and sparsity by projecting it onto the bird's-eye view and realizes real-time 3D object detection based on LiDAR. However, Complex-YOLO has no object height detection, a shallow network depth, and poor small-size object detection accuracy. To address these issues, this paper has made the following improvements: (1) adds a multi-scale feature fusion network to improve the algorithm's capability to detect small-size objects; (2) uses a more advanced RepVGG as the backbone network to improve network depth and overall detection performance; and (3) adds an effective height detector to the network to improve the height detection. Through experiments, we found that our algorithm's accuracy achieved good performance on the KITTI dataset, while the detection speed and memory usage were very superior, 48FPS on RTX3070Ti and 20FPS on GTX1060, with a memory usage of 841Mib.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1662-5218  
入藏号:  
WOS:000941717300001  
PubMed ID:  
36876303  
Indexed Date:  
2023-03-24  
Institution Address:  
China Jiliang Univ, Coll Mech & Elect Engn, Hangzhou, Peoples R China  
Institution:

### 820. License Plate Location Detection Algorithm Based on Improved YOLOv3 in

摘要: Aiming at the problem of the difficulty of license plate positioning, slow detection speed and low detection accuracy in complex scenes such as lighting, multi-vehicle and low resolution, an improved method based on YOLOv3 is proposed. Firstly, the label information of the example is clustered by K-means++ method to obtain a new anchor size. And then, the improved thin feature extraction network(DarkNet41) is used to improve the detection efficiency of the model and reduce computational consumption. Moreover, multi-scale feature fusion is improved from 3-scale prediction to 4-scale prediction and improved Inception-SE structure is added to the detection network to improve the accuracy of detection. Finally, CIoU is selected as a loss function. The data is enhanced with the Multi-Scale Retinex(MSR) algorithm. Experimental analysis shows that the improved algorithm's mAP reaches 98.84% and the detection speed reaches 36.4 frame/s, which has better accuracy and real-time performance compared with the YOLOv3 model and other algorithms.  
摘要:  
针对在光照、多车辆和低分辨率等复杂场景下车牌定位困难、检测速度慢和精度低等问题,提出了一种改进YOLOv3的方法。采用K-means++方法对实例的标签信息进行聚类分析获取新的anchor尺寸,通过改进后的精简特征提取网络(DarkNet41)来提高模型的检测效率并降低计算消耗。此外,改进了多尺度特征融合,由3尺度预测增加至4尺度预测并在检测网络中加入了改进后的Inception-SE结构来提高检测的精度,选取了CIoU作为损失函数。预处理方面用MSR(Multi-Scale Retinex)算法对数据进行增强。实验分析表明,采用该算法mAP(均值平均精度)达到了98.84%,检测速度达到36.4帧/s,与YOLOv3模型以及其他算法相比具有更好的准确性和实时性。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
7  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
8  
ISSN:  
1002-8331  
入藏号:  
CSCD:6939742  
Indexed Date:  
2021-07-02  
Institution Address:  
中北大学软件学院;;山西省军民融合软件技术工程研究中心, ;;山西省军民融合软件技术工程研究中心, 太原;;太原, ;; 030051;;030051  
Institution Address:  
School of Software, North China University;;Shanxi Military-Civilian Integration Software Technology Engineering Research Center, ;;Shanxi Military-Civilian Integration Software Technology Engineering Research Center, Taiyuan;;Taiyuan, ;; 030051;;030051  
Institution:  
中北大学软件学院;;山西省军民融合软件技术工程研究中心  
Institution:  
School of Software, North China University;;Shanxi Military-Civilian Integration Software Technology Engineering Research Center

### 821. Large scale instance segmentation of outdoor environment based on

摘要: Instance segmentation is a challenging task that requires both instance-level and pixel-level prediction and it has a wide range of applications in autonomous driving, video analysis, scene understandingand so on. The currently dominant instance segmentation methods have excellent accuracy, but they are slow, and the processing speed will be even less satisfactory if the input is a large-scale image. In order to improve the efficiency and accuracy of instance segmentation of large-scale images, this article modifies the backbone network based on YOLACT network, adds a multi-information fusion module and provides an improved BiFPN method to achieve multi-scale feature fusion, while adding two branches to the first level detector RetinaNet to achieve instance segmentation. The network model is tested on Cityscapes dataset and the results of the experiments show that the improved instance segmentation network in this article improves the accuracy while ensuring the speed of segmentation. The optimized network model size was reduced by 17% compared to YOLACT, and the mAP, mAP50, and mAP75 were improved by 18.3%, 32.1%, and 24.6%, respectively.  
在 BIOSIS Citation Index 中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1532-0626  
eISSN:  
1532-0634  
入藏号:  
WOS:000863957400001  
Indexed Date:  
2022-10-15  
Institution Address:  
Wuhan Univ Sci & Technol, Minist Educ, Key Lab Met Equipment & Control Technol, Wuhan 430081, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Res Ctr Biomimet Robot & Intelligent Measurement, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Hubei Key Lab Mech Transmiss & Mfg Engn, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Precis Mfg Res Inst, Wuhan, Peoples R China  
Institution Address:  
Hubei Longzhong Lab, Xiangyang, Peoples R China  
Institution Address:  
China Three Gorges Univ, Hubei Key Lab Hydroelect Machinery Design & Maint, Yichang, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Hubei Longzhong Lab  
Institution:

### 822. Radio frequency interference identification using dual cross-attention

摘要: Radio astronomy plays a very important role in promoting scientific progress and unraveling the mysteries of the universe. However, radio telescopes are inevitably affected by radio frequency interference (RFI) when receiving radio signals, which leads to a reduction in data quality and has a serious impact on the formation of correct scientific conclusions. Therefore, it is essential to identify the RFI present in the observational data. In order to effectively identify RFI, improve the existing RFI identification methods that suffer from missed detections, and enhance the performance of RFI identification, this paper proposes a novel method that combines a dual cross-attention mechanism with multi-scale feature fusion. Experimental studies were conducted using the observational data from the 40-meter radio telescope at the Yunnan Astronomical Observatory of the Chinese Academy of Sciences. The proposed method achieved scores of 92.49%, 83.90%, and 87.99% in terms of precision, recall and F1-score, respectively. It outperformed existing methods (U-Net, RFI-Net, R-Net6, RFI-GAN, EMSCA-UNet) in recall and F1-score, effectively reducing the occurrence of missed detections and improving the overall performance of radio frequency interference identification.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2213-1337  
eISSN:  
2213-1345  
入藏号:  
WOS:001341133100001  
Indexed Date:  
2024-10-31  
Institution Address:  
Kunming Univ Sci & Technol, Fac Informat Engn & Automat, 727 Jingming South Rd, Kunming 650500, Yunnan, Peoples R China  
Institution Address:  
Yunnan Key Lab Comp Technol Applicat, 727 Jingming South Rd, Kunming 650500, Yunnan, Peoples R China  
Institution Address:  
Chinese Acad Sci, Yunnan Observ, 396 Yangfangwang, Kunming 650000, Yunnan, Peoples R China  
Institution:  
Institution:  
Yunnan Key Lab Comp Technol Applicat  
Institution:

### 823. MTSDet: multi-scale traffic sign detection with attention and path

摘要: To solve the problem that existing traffic signs are not easily detected leading to low detection performance due to their small sizes and external factors such as weather conditions, this paper proposes a traffic sign detection method, MTSDet (Multi-scale Traffic Sign Detection with attention and path aggregation), which focuses on the multi-scale detection problem and effectively improves the detection performance. First, the method efficiently extracts semantic features by introducing the Attention Mechanism Network(AMNet), and then feeds the multi-scale semantic features into Path Aggregation Feature Pyramid Network(PAFPN) for multi-scale feature fusion to obtain multi-scale advanced semantic features. Finally, the multi-scale advanced semantic feature map is deformable interest pooled to effectively enhance the multi-scale object detection modeling capability. In this paper, the above method is validated by two classical datasets, German traffic sign detection dataset and Chinese traffic sign detection dataset, which achieve 92.9% and 94.3% mAP, respectively, and have obvious detection accuracy improvement when compared with other classical advanced algorithms, effectively proving the superiority and generalization of the algorithm in this paper. Code is available at haps://github.com/why529913/MTSDet  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
22  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
被引频次合计:  
25  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000782550400001  
Indexed Date:  
2022-04-25  
Institution Address:  
Xinjiang Univ, Software Coll, Urumqi 830000, Peoples R China  
Institution Address:  
Key Lab Signal Detect & Proc Xinjiang Uygur Auton, Urumqi 830000, Peoples R China  
Institution Address:  
Xinjiang Univ, Key Lab Software Engn, Urumqi 830000, Peoples R China  
Institution:  
Institution:  
Key Lab Signal Detect & Proc Xinjiang Uygur Auton  
Institution:

### 824. Multi-Feature Information Complementary Detector: A High-Precision

摘要: Remote sensing for image object detection has numerous important applications. However, complex backgrounds and large object-scale differences pose considerable challenges in the detection task. To overcome these issues, we proposed a one-stage remote sensing image object detection model: a multi-feature information complementary detector (MFICDet). This detector contains a positive and negative feature guidance module (PNFG) and a global feature information complementary module (GFIC). Specifically, the PNFG is used to refine features that are beneficial for object detection and explore the noisy features in a complex background of abstract features. The proportion of beneficial features in the feature information stream is increased by suppressing noisy features. The GFIC uses pooling to compress the deep abstract features and improve the model's ability to resist feature displacement and rotation. The pooling operation has the disadvantage of losing detailed feature information; thus, dilated convolution is introduced for feature complementation. Dilated convolution increases the receptive field of the model while maintaining an unchanged spatial resolution. This can improve the ability of the model to recognize long-distance dependent information and establish spatial location relationships between features. The detector proposed also improves the detection performance of objects at different scales in the same image using a dual multi-scale feature fusion strategy. Finally, classification and regression tasks are decoupled in space using a decoupled head. We experimented on the DIOR and NWPU VHR-10 datasets to demonstrate that the newly proposed MFICDet achieves competitive performance compared to current state-of-the-art detectors.  
Web of Science 核心合集中的 "被引频次":  
7  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
eISSN:  
2072-4292  
入藏号:  
WOS:000856779100001  
Indexed Date:  
2022-09-28  
Institution Address:  
PLA Strateg Support Force Informat Engn Univ, Inst Geospatial Informat, Zhengzhou 450001, Peoples R China  
Institution:

### 825. Optical Remote Sensing Ship Recognition and Classification Based on

摘要: Due to the special characteristics of the shooting distance and angle of remote sensing satellites, the pixel area of ship targets is small, and the feature expression is insufficient, which leads to unsatisfactory ship detection performance and even situations such as missed and false detection. To solve these problems, this paper proposes an improved-YOLOv5 algorithm mainly including: (1) Add the Convolutional Block Attention Module (CBAM) into the Backbone to enhance the extraction of target-adaptive optimal features; (2) Introduce a cross-layer connection channel and lightweight GSConv structures into the Neck to achieve higher-level multi-scale feature fusion and reduce the number of model parameters; (3) Use the Wise-IoU loss function to calculate the localization loss in the Output, and assign reasonable gradient gains to cope with differences in image quality. In addition, during the preprocessing stage of experimental data, a median+bilateral filter method was used to reduce interference from ripples and waves and highlight the information of ship features. The experimental results show that Improved-YOLOv5 has a significant improvement in recognition accuracy compared to various mainstream target detection algorithms; compared to the original YOLOv5s, the mean Average Precision (mAP) improved by 3.2% and the Frames Per Second (FPN) accelerated by 8.7%.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
eISSN:  
2072-4292  
入藏号:  
WOS:001119173100001  
Indexed Date:  
2023-12-21  
Institution Address:  
Dalian Maritime Univ, Nav Coll, Dalian 116026, Peoples R China  
Institution Address:  
Key Lab Nav Safety Guarantee, Dalian 116026, Peoples R China  
Institution:  
Institution:  
Key Lab Nav Safety Guarantee

### 826. GD-PAN: a multiscale fusion architecture applied to object detection in

摘要: Because unmanned aerial vehicle aerial images are captured at high altitude, they contain objects at different viewpoints and in different states, resulting in two major problems in the images: target size imbalance and category semantic imbalance. To alleviate the impact of the above problems on object detection in unmanned aerial vehicle aerial images, this paper integrates the classic path aggregation network with the concept of gather-and-distribute mechanism, proposing a novel feature fusion architecture termed as gather-and-distribute path aggregation network(GD-PAN). GD-PAN leverages the advantages of both fusion approaches, resulting in a more thorough and comprehensive fusion of multi-scale feature maps, dramatically improving the refinement of the network for small target objects. Additionally, we introduce wider coordinate attention to optimize generation of coordinate offset and modulation scalar in deformable ConvNets version 2, which is applied to the basic network module of GD-PAN to enhance the feature extraction and fusion capability of the network for objects in different shapes and states. GD-PAN is a general improvement scheme that can be migrated to object detection algorithms based on the structure of path aggregation networks to improve their fusion ability on multi-scale features as well as small object detection performance. On the VisDrone dataset, the proposed method improves AP by 3.2%, mAP by 2.2%, and AP-small by 2.7% when used for real-time DETR, and when applied to YOLOv7 and YOLOv8, AP-small improves by 2% and 3.3%, respectively.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
0942-4962  
eISSN:  
1432-1882  
入藏号:  
WOS:001216152400001  
Indexed Date:  
2024-05-13  
Institution Address:  
Beijing Union Univ, Coll Smart City, Beijing 100101, Peoples R China  
Institution:

### 827. MAD-UNet: A Multi-Region UAV Remote Sensing Network for Rural Building

摘要: For the development of an idyllic rural landscape, an accurate survey of rural buildings is essential. The extraction of rural structures from unmanned aerial vehicle (UAV) remote sensing imagery is prone to errors such as misclassifications, omissions, and subpar edge detailing. This study introduces a multi-scale fusion and detail enhancement network for rural building extraction, termed the Multi-Attention-Detail U-shaped Network (MAD-UNet). Initially, an atrous convolutional pyramid pooling module is integrated between the encoder and decoder to enhance the main network's ability to identify buildings of varying sizes, thereby reducing omissions. Additionally, a Multi-scale Feature Fusion Module (MFFM) is constructed within the decoder, utilizing superficial detail features to refine the layered detail information, which improves the extraction of small-sized structures and their edges. A coordination attention mechanism and deep supervision modules are simultaneously incorporated to minimize misclassifications. MAD-UNet has been tested on a private UAV building dataset and the publicly available Wuhan University (WHU) Building Dataset and benchmarked against models such as U-Net, PSPNet, DeepLabV3+, HRNet, ISANet, and AGSCNet, achieving Intersection over Union (IoU) scores of 77.43% and 91.02%, respectively. The results demonstrate its effectiveness in extracting rural buildings from UAV remote sensing images across different regions.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
5  
eISSN:  
1424-8220  
入藏号:  
WOS:001210886800001  
PubMed ID:  
38676009  
Indexed Date:  
2024-05-10  
Institution Address:  
North China Inst Aerosp Engn, Sch Remote Sensing & Informat Engn, Langfang 065000, Peoples R China  
Institution Address:  
Hebei Collaborat Innovat Ctr Space Remote Sensing, Langfang 065000, Peoples R China  
Institution:  
Institution:  
Hebei Collaborat Innovat Ctr Space Remote Sensing

### 828. Advanced Lung Disease Detection: CBAM-Augmented, Lightweight

摘要: Introduction This paper presents a multichannel deep-learning method for detecting lung diseases using chest X-ray images. Using EfficientNetB0 through EfficientNetB7 pretrained models, the methodology offers improved performance in classifying COVID-19, viral pneumonia, and normal chest X-rays.Methods The EfficientNetB2 model was customized by incorporating Squeeze-and-Excitation (SE) blocks and the Convolutional Block Attention Module (CBAM) to improve the model's attention mechanisms. Additional convolutional layers were added for improved feature extraction, and multi-scale feature fusion was implemented to capture features at different scales.Results In this study, 99.3% of the unseen chest X-ray images were identified using the proposed model. It demonstrated superior performance, surpassing existing techniques and highlighting its robustness and generalizability on unseen data samples.Conclusion Moreover, visualization techniques were used to inspect the intermediate layers of the model, providing deeper insights into its processing and interpretation of medical images. The proposed method offers healthcare radiologists a valuable tool for rapid and accurate point of care diagnoses.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1573-4056  
eISSN:  
1875-6603  
入藏号:  
WOS:001448483300001  
PubMed ID:  
39568109  
Indexed Date:  
2025-03-26  
Institution Address:  
Vellore Inst Technol, Sch Comp Sci & Engn, Chennai, India  
Institution:

### 829. SDSNet: Building Extraction in High-Resolution Remote Sensing Images

摘要: Building extraction refers to the automatic identification and separation of buildings from the background in remote sensing images. It plays a significant role in urban planning, land management, and disaster monitoring. Deep-learning methods have shown advantages in building extraction, but they still face challenges such as variations in building types, object occlusions, and complex backgrounds. To address these issues, SDSNet, a deep convolutional network that incorporates global multi-scale feature extraction and cross-level feature fusion, is proposed. SDSNet consists of three modules: semantic information extraction (SIE), multi-level merge (MLM), and semantic information fusion (SIF). The SIE module extracts contextual information and improves recognition of multi-scale buildings. The MLM module filters irrelevant details guided by high-level semantic information, aiding in the restoration of edge details for buildings. The SIF module combines filtered detail information with extracted semantic information for refined building extraction. A series of experiments conducted on two distinct public datasets for building extraction consistently demonstrate that SDSNet outperforms the state-of-the-art deep-learning models for building extraction tasks. On the WHU building dataset, the overall accuracy (OA) and intersection over union (IoU) achieved impressive scores of 98.86% and 90.17%, respectively. Meanwhile, on the Massachusetts dataset, SDSNet achieved OA and IoU scores of 94.05% and 71.6%, respectively. SDSNet exhibits a unique advantage in recovering fine details along building edges, enabling automated and intelligent building extraction. This capability effectively supports urban planning, resource management, and disaster monitoring.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
6  
eISSN:  
2072-4292  
入藏号:  
WOS:001140618300001  
Indexed Date:  
2024-02-02  
Institution Address:  
China Univ Geosci, Sch Comp Sci, Wuhan 430078, Peoples R China  
Institution Address:  
China Univ Geosci, Hubei Key Lab Geol Survey & Evaluat, Minist Educ, Wuhan 430078, Peoples R China  
Institution Address:  
China Geol Survey, Xining Ctr Nat Resources Comprehens Survey, Xining 810000, Peoples R China  
Institution Address:  
China Univ Geosci, State Key Lab Geol Proc & Mineral Resources, Wuhan 430078, Peoples R China  
Institution Address:  
China Univ Geosci, Key Lab Intelligent Geoinformat Proc, Wuhan 430078, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 830. Blind Detection of Broadband Signal Based on Weighted Bi-Directional

摘要: With the development of wireless technology, signals propagating in space are easy to mix, so blind detection of communication signals has become a very practical and challenging problem. In this paper, we propose a blind detection method for broadband signals based on a weighted bi-directional feature pyramid network (BiFPN). The method can quickly perform detection and automatic modulation identification (AMC) on time-domain aliased signals in broadband data. Firstly, the method performs a time-frequency analysis on the received signals and extracts the normalized time-frequency images and the corresponding labels by short-time Fourier transform (STFT). Secondly, we build a target detection model based on YOLOv5 for time-domain mixed signals in broadband data and learn the features of the time-frequency distribution image dataset of broadband signals, which achieves the purpose of training the model. The main improvements of the algorithm are as follows: (1) a weighted bi-directional feature pyramid network is used to achieve a simple and fast multi-scale feature fusion approach to improve the detection probability; (2) the Efficient-Intersection over Union (EIOU) loss function is introduced to achieve high accuracy signal detection in a low Signal-Noise Ratio (SNR) environment. Finally, the time-frequency images are detected by an improved deep network model to complete the blind detection of time-domain mixed signals. The simulation results show that the method can effectively detect the continuous and burst signals in the broadband communication signal data and identify their modulation types.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:000930073600001  
PubMed ID:  
36772564  
Indexed Date:  
2023-03-09  
Institution Address:  
Fuzhou Univ, Sch Adv Mfg, Fujian 362200, Peoples R China  
Institution Address:  
Fuzhou Univ, Coll Phys & Informat Engn, Fujian 350108, Peoples R China  
Institution:  
Institution:

### 831. Multi-Site Air Quality Index Forecasting Based on Spatiotemporal

摘要: Efficient and accurate air quality forecasting contributes significantly to environmental governance, health promotion, and the development of smart cities. However, few existing models can achieve multi-scale feature fusion and long sequence time series modeling in the prediction process. This study proposes a PatchTST-Enhanced model for multi-site air quality forecasting based on spatiotemporal distribution. It uses daily air quality data from 11 prefecture-level cities in Hebei Province from December 2, 2013, to October 12, 2023, to train the model. The new model demonstrates robust performance in Hebei's air quality forecasting (PreLen =96: MSE =0.5408, MAE =0.5408, RSE =0.5408; PreLen =192: MSE =0.2795, MAE =0.2795, RSE =0.2795; PreLen =336: MSE =0.6779, MAE =0.6779, RSE =0.6779; PreLen =720: MSE =0.6779, MAE =0.6779, RSE =0.6779), and the prediction accuracy has been significantly improved compared to both the pre-optimization model and other existing models. The PatchTST-Enhanced outperforms the PatchTST and improves it through four optimization modules: CGAttention, SiLU activation, AdamW optimizer, and SmoothL1 Loss function. By incorporating spatiotemporal features, the PatchTST-Enhanced can address the challenge of combining spatial and large temporal scales in air quality forecasting. The results provide critical information to protect health and improve the environment for the public.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001320523400001  
Indexed Date:  
2024-10-01  
Institution Address:  
Hebei GEO Univ, Sch Econ, Shijiazhuang 050031, Peoples R China  
Institution Address:  
Hebei GEO Univ, Hebei Ctr Ecol & Environm Geol Res, Shijiazhuang 050031, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, Sch Foreign Languages, Chengdu 610031, Peoples R China  
Institution:  
Institution:  
Institution:

### 832. Remu-Net: Multi-Branch Net Framework for 3D Brain Tumor Segmentation

摘要: As one of the fatal human diseases, early detection of brain tumors can effectively save patients' lives. Brain tumor image segmentation is of great practical importance for physicians to perform brain tumor diagnoses quickly. Due to the data complexity of 3D brain images, it is impractical to segment out tumor regions manually, so automatic and reliable methods can be utilized instead of manual work to achieve accurate segmentation of tumor regions. In this paper, we propose an end-to-end, more efficient brain tumor MRI segmentation model, REMU-Net, for the problems of multi-scale feature extraction and difficulty in small target feature extraction in 3D brain tumor image segmentation. Firstly, design and use the multi-channel parallel M-RepVGG module as a decoder to achieve multi-scale feature fusion. Secondly, embedding dilated convolution with different dilated rates in the DM-RepVGG module of the encoder to better extract features at different scales. Finally, introduce the expectation-maximizing attention in the network to better extract the features of the internal details of the tumor. The experimental results on the BraTS2018 validation dataset are Dice scores of 80.93%, 90.13%, and 86.15%, respectively. Experimental results on the BraTS2019 validation dataset can be achieved with Dice scores of 78.29%, 90.65%, and 82.77%, respectively.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0218-1266  
eISSN:  
1793-6454  
入藏号:  
WOS:000838057100016  
Indexed Date:  
2022-08-23  
Institution Address:  
Dalian Univ, Coll Informat Engn, Dalian, Liaoning, Peoples R China  
Institution Address:  
Dalian Univ, Key Lab Adv Design & Intelligent Comp, Minist Educ, Dalian, Liaoning, Peoples R China  
Institution Address:  
Hainan Univ, Coll Comp Sci & Technol, Haikou, Hainan, Peoples R China  
Institution:  
Institution:  
Institution:

### 833. FEFR: Feature early fusion and reconstruction for online knowledge

摘要: Online knowledge distillation breaks the pre-determined strong and weak teacher-student models, it provides a new way of thinking about knowledge distillation. However, the current online methods often use the Logits-based prediction distribution, and the features containing rich semantic information are rarely used. Even if the feature-based methods are used, they only operate on the last layer of the network, without further exploring the representation knowledge of the middle layer feature map. To address the above issues, we propose an innovative feature early fusion and reconstruction (FEFR) method for online knowledge distillation which entails four essential components: multi-scale feature extraction and intermediate layer feature early fusion, reconstruction of features, dual-attention and overall fusion module in this paper. We propose early fusion by "sum" operation for feature matrices between different layers and advance fusion to improve the feature map representation. In order to enhance the communication ability between groups to obtain features, the features were reconstructed. We create a dual-attention to enhance the critical channel and spatial regions adaptively in order to collect more accurate information. The previously processed feature maps are combined and fused using feature fusion, which also aids in student models training. A study of the network architectures of CIFAR-10, CIFAR-100, CINIC-10 and ImageNet 2012 shows that FEFR provides more useful characterization knowledge for refinement and improves accuracy by about 0.5% compared to other methods.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1064-1246  
eISSN:  
1875-8967  
入藏号:  
WOS:001120921100019  
Indexed Date:  
2024-02-02  
Institution Address:  
Anhui Univ Sci & Technol, Sch Comp Sci & Engn, Huainan 232001, Peoples R China  
Institution Address:  
Anhui Univ Sci & Technol, Inst Environm Friendly Mat & Occupat Hlth, Wuhu, Peoples R China  
Institution Address:  
Jinling Inst Technol, Jiangsu Al Transportat Innovat & Applicat Engn Re, Nanjing, Peoples R China  
Institution Address:  
Tongji Univ, Coll Elect & Informat Engn, Shanghai, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 834. FiFoNet: Fine-Grained Target Focusing Network for Object Detection in

摘要: Detecting objects from images captured by Unmanned Aerial Vehicles (UAVs) is a highly demanding task. It is also considered a very challenging task due to the typically cluttered background and diverse dimensions of the foreground targets, especially small object areas that contain only very limited information. Multi-scale representation learning presents a remarkable approach to recognizing small objects. However, this strategy ignores the combination of the sub-parts in an object and also suffers from the background interference in the feature fusion process. To this end, we propose a Fine-grained Target Focusing Network (FiFoNet) which can effectively select a combination of multi-scale features for an object and block background interference, which further revitalizes the differentiability of the multi-scale feature representation. Furthermore, we propose a Global-Local Context Collector (GLCC) to extract global and local contextual information and enhance low-quality representations of small objects. We evaluate the performance of the proposed FiFoNet on the challenging task of object detection in UAV images. A comparison of the experiment results on three datasets, namely VisDrone2019, UAVDT, and our VisDrone\_Foggy, demonstrates the effectiveness of FiFoNet, which outperforms the ten baseline and state-of-the-art models with remarkable performance improvements. When deployed on an edge device NVIDIA JETSON XAVIER NX, our FiFoNet only takes about 80 milliseconds to process an drone-captured image.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
13  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
15  
eISSN:  
2072-4292  
入藏号:  
WOS:000845304800001  
Indexed Date:  
2022-09-03  
Institution Address:  
Xidian Univ, Guangzhou Inst Technol, Guangzhou 510555, Peoples R China  
Institution Address:  
Univ Technol Sydney, Global Big Data Technol Ctr, Ultimo, NSW 2007, Australia  
Institution Address:  
Xidian Univ, Sch Comp Sci & Technol, Xian 710071, Peoples R China  
Institution Address:  
Tsinghua Univ, Dept Elect Engn, Beijing 100084, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Sch Intelligent Syst Engn, Shenzhen 518107, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 835. Lung-YOLO: Multiscale feature fusion attention and cross-layer

摘要: Objective: Lung cancer is a significant public health problem worldwide, and its mortality and morbidity rates are among the highest of cancers. At the same time, early diagnosis of nodules can significantly improve the survival rate of patients. Therefore, this paper proposes the Lung-YOLO algorithm for lung CT image detection based on YOLOv6. Methods: First, to enable the network to detect nodules of different sizes and to minimize missed detections, we introduce the Multi-scale Dual-branch Attention (MSDA) mechanism in the feature extraction part of the network. The input features undergo continual dilation convolutions, effectively establish remote dependencies, and the fused multiscale contextual information expands the receptive field while enhancing the model's ability to detect targets of different sizes, which not only contains precise category and location information but also enables the allocation of attentional weights to generate more pixel-level attention. Then, the fused features access the dual-branch attention module to shift the model's attention to the target nodules, and the composed dual-branch structure captures cross-dimensional interactions and realizes inter-dimensional dependencies, effectively improving the detection performance. Second, during the feature transfer process, the original Bidirectional Feature Pyramid structure (RepBiFPAN) suffers from the loss of detailed information such as texture and color, making it challenging to localize target nodules accurately. To address this, we propose the Cross-layer Aggregation Module (CLAM), by cross-layer aggregating the multi-level feature layer of the backbone with the multi-level detection layer of the head, which preserves the multi-level fine-grained information that may be lost during the feature transfer process, which is crucial for the detection of small targets. Finally, the module proposed in this paper can be easily incorporated into any detection framework for plug-and-play. Results: Our method achieves accuracy, precision, recall, and mAP of 97.5 %, 96.5 %, 96.9 %, and 97.9 % on the LUNA16 dataset, and 95.1 %, 94.3 %, 93.4 %, and 95.9 % on the LIDC-IDRI dataset, respectively, surpassing many existing state-of-the-art detection methods. Moreover, the inference speed is 22.8 ms and 28.4 ms per image with 30.6 M parameters, respectively.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:001314843800001  
Indexed Date:  
2024-09-23  
Institution Address:  
Henan Polytech Univ, Sch Comp Sci & Technol, Jiaozuo 454000, Henan, Peoples R China  
Institution Address:  
Univ Leicester, Sch Comp & Math Sci, Leicester LE1, England  
Institution Address:  
King Abdulaziz Univ, Fac Comp & Informat Technol, Dept Informat Technol, Jeddah 21589, Saudi Arabia  
Institution:  
Institution:  
Institution:

### 836. Boundary-aware small object detection with attention and interaction

摘要: Object detection is a critical technology for the intelligent analytical processing of images captured by drones. The objects usually come in various scales and can be extremely small. Existing detection methods are inherently based on pyramid hierarchy architectures to extract multi-scale features and provide better feature representation for small objects. Nevertheless, they inevitably dilute the representation of details in low-level features during top-down feature fusion and are totally unconcerned with whether the fused feature fits the objects of specific scales within a layer. Moreover, the pyramid can only implicitly fuse the spatial context, which makes the fused features cannot receive fine spatial location information for object localization. In this work, we propose an effective boundary-aware network with attention refinement and spatial interaction to tackle the above challenges. Specifically, we first present a highly effective yet simple boundary-aware detection head (BAH), which directly guides representation learning of object structure semantics in the prediction layer to preserve object-related boundary semantics. Additionally, the attentional feature parallel fusion (AFPF) module offers multi-scale feature encoding capability in a parallel triple fusion fashion and adaptively selects features appropriate for objects of certain scales. Furthermore, we design a spatial interactive module (SIM) to preserve fine spatial detail through cross-spatial feature association. Extensive experiments prove that the proposed network significantly outperforms the state-of-the-art methods, in which we achieve 33.1 mAP and 56.5 AP50 on the VisDrone benchmark, 63.4 mAP and 94 AP50 on the NWPU VHR-10 benchmark. The source code will be released.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001100186900001  
Indexed Date:  
2023-12-07  
Institution Address:  
China Univ Min & Technol, Sch Comp Sci & Technol, Xuzhou 221116, Peoples R China  
Institution Address:  
Minist Educ Peoples Republ China, Engn Res Ctr Mine Digitizat, Xuzhou 221116, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Dept Comp Sci & Engn, Shanghai 200240, Peoples R China  
Institution:  
Institution:  
Minist Educ Peoples Republ China  
Institution:

### 837. High-Resolution Multi-Scale Feature Fusion Network for Running Posture

摘要: Running posture estimation is a specialized task in human pose estimation that has received relatively little research attention due to the lack of appropriate datasets. To address this issue, this paper presents the construction of a new benchmark dataset called "Running Human", which was specifically designed for running sports. This dataset contains over 1000 images along with comprehensive annotations for 1288 instances of running humans, including bounding boxes and keypoint annotations on the human body. Additionally, a Receptive Field Spatial Pooling (RFSP) module was developed to tackle the challenge of joint occlusion, which is common in running sports images. This module was incorporated into the High-Resolution Network (HRNet) model, resulting in a novel network model named the Running Human Posture Network (RHPNet). By expanding the receptive field and effectively utilizing multi-scale features extracted from the multi-branch network, the RHPNet model significantly enhances the accuracy of running posture estimation. On the Running Human dataset, the proposed method achieved state-of-the-art performance. Furthermore, experiments were conducted on two benchmark datasets. Compared to the state-of-the-art ViTPose-L method, when applied to the COCO dataset, RHPNet demonstrated comparable prediction accuracy while utilizing only one tenth of the parameters and one eighth of the floating-point operations (FLOPs). On the MPII dataset, RHPNet achieves a PCKh@0.5 score of 92.0, which is only 0.5 points lower than the state-of-the-art method, PCT. These experimental results provide strong validation for the effectiveness and excellent generalization ability of the proposed method.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001201187800001  
Indexed Date:  
2024-04-18  
Institution Address:  
Yunnan Normal Univ, Sch Informat Sci & Technol, Kunming 650500, Peoples R China  
Institution:

### 838. IBFusion: An Infrared and Visible Image Fusion Method Based on Infrared

摘要: The fusion of infrared (IR) and visible (VIS) images aims to capture complementary information from diverse sensors, resulting in a fused image that enhances the overall human perception of the scene. However, existing fusion methods face challenges preserving diverse feature information, leading to cross-modal interference, feature degradation, and detail loss in the fused image. To solve the above problems, this paper proposes an image fusion method based on the infrared target mask and bimodal feature extraction strategy, termed IBFusion. Firstly, we define an infrared target mask, employing it to retain crucial information from the source images in the fused result. Additionally, we devise a mixed loss function, encompassing content loss, gradient loss, and structure loss, to ensure the coherence of the fused image with the IR and VIS images. Then, the mask is introduced into the mixed loss function to guide feature extraction and unsupervised network optimization. Secondly, we create a bimodal feature extraction strategy and construct a Dual-channel Multi-scale Feature Extraction Module (DMFEM) to extract thermal target information from the IR image and background texture information from the VIS image. This module retains the complementary information of the two source images. Finally, we use the Feature Fusion Module (FFM) to fuse the features effectively, generating the fusion result. Experiments on three public datasets demonstrate that the fusion results of our method have prominent infrared targets and clear texture details. Both subjective and objective assessments are better than the other twelve advanced algorithms, proving our method's effectiveness.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1520-9210  
eISSN:  
1941-0077  
入藏号:  
WOS:001358607300010  
Indexed Date:  
2024-12-09  
Institution Address:  
Yanshan Univ, Sch Informat Sci & Engn, Key Lab Special Fiber & Fiber Sensor Hebei Prov, Qinhuangdao 066004, Peoples R China  
Institution Address:  
Beijing Inst Technol, Sch Integrated Circuits & Elect, Beijing 100081, Peoples R China  
Institution Address:  
Beijing Inst Technol, Tangshan Res Inst, Beijing 100081, Peoples R China  
Institution Address:  
Beijing Inst Technol, Coll Informat & Elect, Beijing 100081, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 839. A Multi-Scale Edge Constraint Network for the Fine Extraction of

摘要: Building extraction based on remote sensing images has been widely used in many industries. However, state-of-the-art methods produce an incomplete segmentation of buildings owing to unstable multi-scale context aggregation and a lack of consideration of semantic boundaries, ultimately resulting in large uncertainties in predictions at building boundaries. In this study, efficient fine building extraction methods were explored, which demonstrated that the rational use of edge features can significantly improve building recognition performance. Herein, a fine building extraction network based on a multi-scale edge constraint (MEC-Net) was proposed, which integrates the multi-scale feature fusion advantages of UNet++ and fuses edge features with other learnable multi-scale features to achieve the effect of prior constraints. Attention was paid to the alleviation of noise interference in the edge features. At the data level, according to the improvement of copy-paste according to the characteristics of remote sensing imaging, a data augmentation method for buildings (build-building) was proposed, which increased the number and diversity of positive samples by simulating the construction of buildings to increase the generalization of MEC-Net. MEC-Net achieved 91.13%, 81.05% and 74.13% IoU on the WHU, Massachusetts and Inria datasets, and it has a good inference efficiency. The experimental results show that MEC-Net outperforms the state-of-the-art methods, demonstrating its superiority. MEC-Net improves the accuracy of building boundaries by rationally using previous edge features.  
在中国科学引文数据库中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
11  
eISSN:  
2072-4292  
入藏号:  
WOS:000941750600001  
Indexed Date:  
2023-03-24  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution Address:  
Hainan Aerosp Informat Res Inst, Key Lab Earth Observat Hainan Prov, Sanya 572029, Peoples R China  
Institution:  
Institution:  
Institution:  
Hainan Aerosp Informat Res Inst

### 840. Space to depth convolution bundled with coordinate attention for

摘要: Surface defects of steel plates unavoidably exist during the industrial production proceeding due to the complex productive technologies and always exhibit some typical characteristics, such as irregular shape, random position, and various size. Therefore, detecting these surface defects with high performance is crucial for producing high-quality products in practice. In this paper, an improved network with high performance based on You Only Look Once version 5 (YOLOv5) is proposed for detecting surface defects of steel plates. Firstly, the Space to Depth Convolution (SPD-Conv) is utilized to make the feature information transforming from space to depth, helpful for preserving the entirety of discriminative feature information to the greatest extent under the proceeding of down-sampling. Subsequently, the coordinate attention mechanism is introduced and embedded into the bottleneck of C3 modules to effectively enhance the weights of some important feature channels, in favor of capturing more important feature information from different channels after SPD-Conv operations. Finally, the Spatial Pyramid Pooling Faster module is replaced by the Spatial Pyramid Pooling Fully Connected Spatial Pyramid Convolution module to further enhance the feature expression capability and efficiently realize the multi-scale feature fusion. The experimental results on NEU-DET dataset show that, compared with YOLOv5, the mAP and mAP50 dramatically increase from 51.7, 87.0 to 61.4, 92.6%, respectively. Meanwhile, the frame rate of 250 FPS implies that it still preserves a well real-time performance. Undoubtedly, the improved algorithm proposed in this paper exhibits outstanding performance, which may be also used to recognize the surface defects of aluminum plates, as well as plastic plates, armor plates and so on in the future.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001194868100002  
Indexed Date:  
2024-04-06  
Institution Address:  
Jiangxi Normal Univ, Coll Phys & Commun Elect, Nanchang 330022, Jiangxi, Peoples R China  
Institution:

### 841. MSFFA-YOLO Network: Multiclass Object Detection for Traffic

摘要: Despite significant progress in vision-based detection methods, the task of detecting traffic objects in foggy weather remains challenging. The presence of fog reduces visibility, which in turn affects the information of traffic objects in videos. However, accurate information regarding the localization and classification of traffic objects is crucial for certain traffic investigations. In this article, we focus on presenting a multiclass object detection method, namely, multiscale feature fusion attention-you only look once (MSFFA-YOLO) network, that can be trained and jointly achieve three tasks: visibility enhancement, object classification, and object localization. In the network, we employ the enhanced YOLOv7 as a detection subnet, which is responsible for learning to locate and classify objects. In the restoration subnet, the MSFFA structure is presented for visibility enhancement. The experimental results on the synthetic foggy datasets show that the presented MSFFA-YOLO can achieve 64.6% accuracy on the FC005 dataset, 67.3% accuracy on the FC01 dataset, and 65.7% accuracy on the FC02 dataset. When evaluated on the natural foggy datasets, the presented MSFFA-YOLO can achieve 84.7% accuracy on the RTTS dataset and 84.1% accuracy on the RW dataset, indicating its ability to accurately detect multiclass traffic objects in real and foggy weather. And, the experimental results show that the presented MSFFA-YOLO can achieve the efficiency of 37 frames per second (FPS). Finally, the experimental results demonstrate the excellent performance of our presented method for object localization and classification in foggy weather. And, when detecting concealed traffic objects in foggy weather, our presented method exhibits superior accuracy. These results substantiate the applicability of our presented method for traffic investigations in foggy weather.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
15  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
15  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001137037600009  
Indexed Date:  
2024-02-02  
Institution Address:  
Southeast Univ, Jiangsu Key Lab Urban ITS, Jiangsu Prov Collaborat Innovat Ctr Modern Urban, Nanjing 211189, Peoples R China  
Institution Address:  
Southeast Univ, Sch Transportat, Nanjing 211189, Peoples R China  
Institution:  
Institution:

### 842. Enhanced Abandoned Object Detection through Adaptive Dual-Background

摘要: Abandoned object detection is a critical task in the field of public safety. However, existing methods perform poorly when detecting small and occluded objects, leading to high false detection and missed detection rates. To address this issue, this paper proposes an abandoned object detection method that integrates an adaptive dual-background model with SAO-YOLO (Small Abandoned Object YOLO). The goal is to reduce false and missed detection rates for small and occluded objects, thereby improving overall detection accuracy. First, the paper introduces an adaptive dual-background model that adjusts according to scene changes, reducing noise interference in the background model. When combined with an improved PFSM (Pixel-based Finite State Machine) model, this enhances detection accuracy and robustness. Next, a network model called SAO-YOLO is designed. Key improvements within this model include the SAO-FPN (Small Abandoned Object FPN) feature extraction network, which fully extracts features of small objects, and a lightweight decoupled head, SODHead (Small Object Detection Head), which precisely extracts local features and enhances detection accuracy through multi-scale feature fusion. Finally, experimental results show that SAO-YOLO increases mAP@0.5 and mAP@0.5:0.95 by 9.0% and 5.1%, respectively, over the baseline model. It outperforms other advanced detection models. Ultimately, after a series of experiments on the ABODA, PETS2006, and AVSS2007 datasets, the proposed method achieved an average detection precious of 91.1%, surpassing other advanced methods. It significantly outperforms other advanced detection methods. This approach notably reduces false and missed detections, especially for small and occluded objects.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
1424-8220  
入藏号:  
WOS:001341625800001  
PubMed ID:  
39460053  
Indexed Date:  
2024-11-02  
Institution Address:  
Shenyang Jianzhu Univ, Coll Comp Sci & Engn, Shenyang 110168, Peoples R China  
Institution Address:  
Liaoning Prov Big Data Management & Anal Lab Urban, Shenyang 110168, Peoples R China  
Institution Address:  
Natl Special Comp Engn Technol Res Ctr, Shenyang Branch, Shenyang 110168, Peoples R China  
Institution:  
Institution:  
Liaoning Prov Big Data Management & Anal Lab Urban  
Institution:  
Natl Special Comp Engn Technol Res Ctr

### 843. Camouflaged object segmentation based on edge enhancement and feature

摘要: The task of camouflaged object segmentation is to accurately classify and localize objects that are highly similar to the background using pixel-level segmentation masks, which is more challenging than traditional object segmentation tasks. Aiming at the problems that the target is highly similar to the surrounding environment, the boundary is blurred, and the contrast is low, a camouflaged target segmentation method based on edge enhancement and feature fusion is constructed. First, a set of edge extraction modules is designed, aiming to accurately segment valid edge priors. Afterwards, a multi-scale feature enhancement module and a cross-level feature aggregation module are introduced to mine multi-scale contextual information within and between layers, respectively. In addition, a simple inter-layer attention module is proposed to effectively filter out the interference information existing after fusion by utilizing the difference between adjacent layers. Finally, accurate prediction results are obtained by combining feature maps of all levels with edge priors step by step. Experimental results show that the model outperforms other algorithms on four camouflaged target benchmark datasets. Among them, the weighted F value increased by 2. 4%, the average absolute error decreased by 7. 2%, and the segmentation speed reached 44. 2 FPS under the RTX 2080Ti hardware environment. Compared with existing methods, this algorithm can segment camouflage targets more accurately.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1007-2780  
eISSN:  
2097-3217  
入藏号:  
WOS:001157493300008  
Indexed Date:  
2024-02-16  
Institution Address:  
Changchun Inst Opt Fine Mech & Phys, Chinese Acad Sci, Changchun 130033, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 844. 被撤回的出版物: Evaluation of Fundus Blood Flow Perfusion in Patients with

摘要: In this paper, we have carefully investigated the clinical phenotype and genotype of patients with Johanson-Blizzard syndrome (JBS) with diabetes mellitus as the main manifestation. Retinal vessel segmentation is an important tool for the detection of many eye diseases and plays an important role in the automated screening system for retinal diseases. A segmentation algorithm based on a multiscale attentional resolution network is proposed to address the problem of insufficient segmentation of small vessels and pathological missegmentation in existing methods. The network is based on the encoder-decoder architecture, and the attention residual block is introduced in the submodule to enhance the feature propagation ability and reduce the impact of uneven illumination and low contrast on the model. The jump connection is added between the encoder and decoder, and the traditional pooling layer is removed to retain sufficient vascular detail information. Two multiscale feature fusion methods, parallel multibranch structure, and spatial pyramid pooling are used to achieve feature extraction under different sensory fields. We collected the clinical data, laboratory tests, and imaging examinations of JBS patients, extracted the genomic DNA of relevant family members, and validated them by whole-exome sequencing and Sanger sequencing. The patient had diabetes mellitus as the main manifestation, with widened eye spacing, low flat nasal root, hypoplastic nasal wing, and low hairline deformities. Genetic testing confirmed the presence of a c.4463 T > C (p.Ile1488Thr) pure missense mutation in the UBR1 gene, which was a novel mutation locus, and pathogenicity analysis indicated that the locus was pathogenic. This patient carries a new UBR1 gene c.4463 T > C pure mutation, which improves the clinical understanding of the clinical phenotypic spectrum of JBS and broadens the genetic spectrum of the UBR1 gene. The experimental results showed that the method achieved 83.26% and 82.56% F1 values on CHASEDB1 and STARE standard sets, respectively, and 83.51% and 81.20% sensitivity, respectively, and its performance was better than the current mainstream methods.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2040-2295  
eISSN:  
2040-2309  
入藏号:  
WOS:000804416600011  
PubMed ID:  
35222891  
Indexed Date:  
2022-06-16  
Institution Address:  
Lianshui Cty Peoples Hosp, Dept Endocrinol, Huaian 223400, Jiangsu, Peoples R China  
Institution Address:  
Xuzhou Med Univ, Huaian Peoples Hosp 2, Dept Ophthalmol, Affiliated Huaian Hosp, Huaian 223001, Jiangsu, Peoples R China  
Institution Address:  
Hongze Dist Peoples Hosp, Dept Ophthalmol, Huaian 223100, Jiangsu, Peoples R China  
Institution:  
Lianshui Cty Peoples Hosp  
Institution:  
Institution:  
Hongze Dist Peoples Hosp

### 845. SparseDet: A Simple and Effective Framework for Fully Sparse LiDAR-Based

摘要: LiDAR-based sparse 3-D object detection plays a crucial role in autonomous driving applications due to its computational efficiency advantages. Existing methods either use the features of a single central voxel as an object proxy or treat an aggregated cluster of foreground points as an object proxy. However, the former cannot aggregate contextual information, resulting in insufficient information expression in object proxies. The latter relies on multistage pipelines and auxiliary tasks, which reduce the inference speed. To maintain the efficiency of the sparse framework while fully aggregating contextual information, in this work, we propose SparseDet that designs sparse queries as object proxies. It introduces two key modules: the local multiscale feature aggregation (LMFA) module and the global feature aggregation (GFA) module, aiming to fully capture the contextual information, thereby enhancing the ability of the proxies to represent objects. The LMFA module achieves feature fusion across different scales for sparse key voxels via coordinate transformations and using nearest neighbor relationships to capture object-level details and local contextual information, whereas the GFA module uses self-attention mechanisms to selectively aggregate the features of the key voxels across the entire scene for capturing scene-level contextual information. Experiments on nuScenes and KITTI demonstrate the effectiveness of our method. Specifically, SparseDet surpasses the previous best sparse detector VoxelNeXt (a typical method using voxels as object proxies) by 2.2% mean average precision (mAP) with 13.5 frames/s on nuScenes and outperforms VoxelNeXt by 1.12% AP(3-D) on hard level tasks with 17.9 frames/s on KITTI. What is more, not only the mAP of SparseDet exceeds that of FSDV2 (a classical method using clusters of foreground points as object proxies) but also its inference speed is 1.3 times faster than FSDV2 on the nuScenes test set. The code has been released in https://github.com/liulin813/SparseDet.git.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001336387700048  
Indexed Date:  
2024-10-29  
Institution Address:  
Beijing Jiaotong Univ, Sch Comp Sci & Technol, Beijing Key Lab Traff Data Anal & Min, Beijing 100044, Peoples R China  
Institution Address:  
Xiamen Univ, Fujian Key Lab Sensing & Comp Smart Cities, Xiamen 361005, Fujian, Peoples R China  
Institution Address:  
Tsinghua Univ, State Key Lab Intelligent Green Vehicle & Mobil, Beijing 100084, Peoples R China  
Institution Address:  
Tsinghua Univ, Sch Vehicle & Mobil, Beijing 100084, Peoples R China  
Institution Address:  
JD Logist, Autonomous Driving Dept X Div, Beijing 101111, Peoples R China  
Institution Address:  
Horizon Robot, Beijing 100190, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
JD Logist  
Institution:  
Horizon Robot

### 846. Compound feature attention network with edge enhancement for low-dose CT

摘要: BACKGROUND: Low-dose CT (LDCT) images usually contain serious noise and artifacts, which weaken the readability of the image.  
OBJECTIVE: To solve this problem, we propose a compound feature attention network with edge enhancement for LDCT denoising (CFAN-Net), which consists of an edge-enhanced module and a proposed compound feature attention block (CFAB).  
METHODS: The edge enhancement module extracts edge details with the trainable Sobel convolution. CFAB consists of an interactive feature learning module (IFLM), a multi-scale feature fusion module (MFFM), and a joint attention module (JAB), which removes noise from LDCT images in a coarse-to-fine manner. First, in IFLM, the noise is initially removed by cross-latitude interactive judgment learning. Second, in MFFM, multi-scale and pixel attention are integrated to explore fine noise removal. Finally, in JAB, we focus on key information, extract useful features, and improve the efficiency of network learning. To construct a high-quality image, we repeat the above operation by cascading CFAB.  
RESULTS: By applying CFAN-Net to process the 2016 NIH AAPM-Mayo LDCT challenge test dataset, experiments show that the peak signal-to-noise ratio value is 33.9692 and the structural similarity value is 0.9198.  
CONCLUSIONS: Compared with several existing LDCT denoising algorithms, CFAN-Net effectively preserves the texture of CT images while removing noise and artifacts.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0895-3996  
eISSN:  
1095-9114  
入藏号:  
WOS:001074432400004  
PubMed ID:  
37355934  
Indexed Date:  
2023-10-22  
Institution Address:  
North Univ China, Sch Informat & Commun Engn, State Key Lab Dynam Testing Technol, 3 Coll Rd, Taiyuan 030051, Shanxi, Peoples R China  
Institution:

### 847. Research on the magnetic resonance imaging brain tumor segmentation

摘要: With the social and economic development and the improvement of people's living standards, smart medical care is booming, and medical image processing is becoming more and more popular in research, of which brain tumor segmentation is an important branch of medical image processing. However, the manual segmentation method of brain tumors requires a lot of time and effort from the doctor and has a great impact on the treatment of patients. In order to solve this problem, we propose a DO-UNet model for magnetic resonance imaging brain tumor image segmentation based on attention mechanism and multi-scale feature fusion to realize fully automatic segmentation of brain tumors. Firstly, we replace the convolution blocks in the original U-Net model with the residual modules to prevent the gradient disappearing. Secondly, the multi-scale feature fusion is added to the skip connection of U-Net to fuse the low-level features and high-level features more effectively. In addition, in the decoding stage, we add an attention mechanism to increase the weight of effective information and avoid information redundancy. Finally, we replace the traditional convolution in the model with DO-Conv to speed up the network training and improve the segmentation accuracy. In order to evaluate the model, we used the BraTS2018, BraTS2019, and BraTS2020 datasets to train the improved model and validate it online, respectively. Experimental results show that the DO-UNet model can effectively improve the accuracy of brain tumor segmentation and has good segmentation performance.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0899-9457  
eISSN:  
1098-1098  
入藏号:  
WOS:000821454800001  
Indexed Date:  
2022-07-12  
Institution Address:  
Chongqing Univ Technol, Sch Artificial Intelligence, Chongqing, Peoples R China  
Institution:

### 848. ResAttenGAN: Simultaneous segmentation of multiple spinal structures on

摘要: An axial MRI image of the lumbar spine generally contains multiple spinal structures and their simultaneous segmentation will help analyze the pathogenesis of the spinal disease, generate the spinal medical report, and make a clinical surgery plan for the treatment of the spinal disease. However, it is still a challenging issue that multiple spinal structures are segmented simultaneously and accurately because of the large diversities of the same spinal structure in intensity, resolution, position, shape, and size, the implicit borders between different structures, and the overfitting problem caused by the insufficient training data. In this paper, we propose a novel network framework ResAttenGAN to address these challenges and achieve the simultaneous and accurate segmentation of disc, neural foramina, thecal sac, and posterior arch. ResAttenGAN comprises three modules, i.e. full feature fusion (FFF) module, residual refinement attention (RRA) module, and adversarial learning (AL) module. The FFF module captures multi-scale feature information and fully fuse the features at all hierarchies for generating the discriminative feature representation. The RRA module is made up of a local position attention block and a residual border refinement block to accurately locate the implicit borders and refine their pixel-wise classification. The AL module smooths and strengthens the higher-order spatial consistency to solve the overfitting problem. Experimental results show that the three integrated modules in ResAttenGAN have advantages in tackling the above challenges and ResAttenGAN outperforms the existing segmentation methods under evaluation metrics.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
15  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
15  
ISSN:  
0933-3657  
eISSN:  
1873-2860  
入藏号:  
WOS:000748991500005  
PubMed ID:  
35115128  
Indexed Date:  
2022-02-06  
Institution Address:  
Beijing Inst Technol, Sch Mech Engn, 5 South Zhongguancun St, Beijing 100081, Peoples R China  
Institution Address:  
Western Univ, Sch Hlth Sci, London, ON N6A 4V2, Canada  
Institution Address:  
Univ Western, Dept Med Imaging & Med Biophys, London, ON N6A 5W9, Canada  
Institution:  
Institution:  
Institution:  
Univ Western

### 849. A Nested UNet Based on Multi-Scale Feature Extraction for Mixed

摘要: Eliminating mixed noise from images is a challenging task because accurately describing the attenuation of noise distribution is difficult. However, most existing algorithms for mixed noise removal solely rely on the local information of the image and neglect the global information, resulting in suboptimal denoising performance when dealing with complex mixed noise. In this paper, we propose a nested UNet based on multi-scale feature extraction (MSNUNet) for mixed noise removal. In MSNUNet, we introduce a U-shaped subnetwork called MSU-Subnet for multi-scale feature extraction. These multi-scale features contain abundant local and global features, aiding the model in estimating noise more accurately and improving its robustness. Furthermore, we introduce a multi-scale feature fusion channel attention module (MSCAM) to effectively aggregate feature information from different scales while preserving intricate image texture details. Our experimental results demonstrate that MSNUNet achieves leading performance in terms of quality metrics and the visual appearance of images.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001061010800001  
Indexed Date:  
2023-09-16  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Software, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Jiangsu Collaborat Innovat Ctr Atmospher Environm, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Normal Univ Special Educ, Coll Math & Informat Sci, Nanjing 210038, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Comp Sci, Nanjing 210044, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 850. MFDiff: multiscale feature diffusion model for segmentation of 3D

摘要: Intracranial aneurysm is a common life-threatening disease, and the rupture of an intracranial aneurysm carries a high risk of morbidity and mortality. Due to their small size in images, it remains a challenging task to accurately extract the intracranial aneurysms in CT images. In this paper, we propose a multi-scale feature diffusion model, named as MFDiff in short, for segmentation of 3D intracranial aneurysm. The proposed MFDiff includes a feature extraction module and a diffusion model. The feature extraction module is designed to extract features of the original image, and the features act as conditional priors to guide the diffusion model to gradually generate segmentation maps. The diffusion model takes a structure similar to U-Net as backbone, and there is a residual multi-scale feature fusion attention module (RMFA) in the diffusion model, which can adapt to intracranial aneurysms of different size due to multi-scale features. A local CT image dataset is employed for experiment, there are both ruptured and unruptured intracranial aneurysms in the images, and the size of intracranial aneurysms is various, even less than 3 mm. Compared with other popular methods, such as U-Net, GLIA-Net, UNETR++ , LinTransUNet, Swin UNETR, the proposed MFDiff shows better performance in intracranial aneurysm segmentation, the segmentation precision is 82.91% when the aneurysms of just size larger than 3 mm are taken into account, and the precision is 75.53% when considering aneurysms of all size.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1433-7541  
eISSN:  
1433-755X  
入藏号:  
WOS:001227402600001  
Indexed Date:  
2024-05-27  
Institution Address:  
Hebei Univ Technol HeBUT, Sch Artificial Intelligence, Tianjin 300401, Peoples R China  
Institution Address:  
Qingdao Univ, Dept Radiol, Affiliated Hosp, Qingdao, Shandong, Peoples R China  
Institution Address:  
Third Cent Hosp Tianjin, Tianjin 300171, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Comp Technol, Beijing 100190, Peoples R China  
Institution:  
Hebei Univ Technol HeBUT  
Institution:  
Institution:  
Third Cent Hosp Tianjin  
Institution:

### 851. Power Line Segmentation Model of Aerial Images Based on Hierarchical

摘要: Automatic power line segmentation is an important prerequisite for the safe operation of intelligent inspection platforms. However, power line segmentation is a small target segmentation problem in complex backgrounds and multiple climatic environments, which is highly prone to encounter false or missed detections. In order to improve the robustness and accuracy of power line segmentation, an end-to-end segmentation model based on hierarchical attention fusion is proposed in combination with an encoderdecoder framework. The model proposes a reduced-dimensional residual convolution unit that increases the network depth while significantly reducing the network parameters, making it easier to deploy in embedded devices, enabling the model to capture global information and emphasize the target regions of powerlines, a chain-based hierarchical attention fusion module is designed for multiscale feature fusion to address the category imbalance problem. To improve the model's attention to the unique line prior features of power lines, the line prior loss function is combined with the Focal loss function and Dice loss function to form a joint loss function to further improve the accuracy of power line segmentation. The experimental results show that the depth of the proposed model network increases to about 2.8 times that of the base network, while the number of parameters is only about 1/3 of the original one. Robust segmentation of power lines can be achieved for both regular weather and foggy weather aerial images. The proposed model can be applied to the field of power inspection, making the inspection more intelligent and efficient.  
摘要:  
电力线自动分割是保证智能检测平台安全运行的重要前提。然而,电力线分割是复杂背景、多种气候环境下的小目标分割问题,极易出现误检、漏检问题。为了提高电力线分割的鲁棒性与准确性,结合编码器-解码器框架,提出了一种基于层级化注意力融合的端到端分割模型。该模型提出一种降维残差卷积单元,增加网络深度的同时大幅度减少网络参数,更易部署于嵌入式设备。为了使模型捕捉到全局信息并强调电力线的目标区域,设计了链式层级化注意力融合模块进行多尺度特征融合,以解决类别不均衡问题。为了提高模型对电力线特有的直线先验特征的关注,提出了直线先验损失函数,并与Focal、Dice损失函数组合成联合损失函数,进一步提高了电力线分割准确度。实验结果表明,所提模型网络深度增加为基础网络的2.8倍左右,而参数量仅为原来的1/3左右。针对常规天气和雾天环境下的航拍图像均能实现电力线的鲁棒分割。所提出模型能应用于电力巡检领域,使巡检更加智能与高效。  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1674-0629  
入藏号:  
CSCD:7779520  
Indexed Date:  
2024-10-12  
Institution Address:  
江苏师范大学电气工程及自动化学院, 徐州, 江苏 221116, 中国  
Institution Address:  
School of Electrical Engineering & Automation, Jiangsu Normal University, Xuzhou, Jiangsu 221116, China  
Institution:  
江苏师范大学电气工程及自动化学院  
Institution:  
School of Electrical Engineering & Automation, Jiangsu Normal University

### 852. FloodKAN: Integrating Kolmogorov-Arnold Networks for Efficient Flood

摘要: Flood events are among the most destructive natural catastrophes worldwide and pose serious threats to socioeconomic systems, ecological environments, and the safety of human life and property. With the advancement of remote sensing technology, synthetic aperture radar (SAR) has provided new means for flood monitoring. However, traditional methods have limitations when dealing with high noise levels and complex terrain backgrounds. To address this issue, in this study, we adopt an improved U-Net model incorporating the Kolmogorov-Arnold Network (KAN), referred to as UKAN, for the efficient extraction of flood inundation extents from multisource remote sensing data. UKAN integrates the efficient nonlinear mapping capabilities of KAN layers with the multiscale feature fusion mechanism of U-Net, enabling better capturing of complex nonlinear relationships and global features. Experiments were conducted on the C2S-MS Floods and MMFlood datasets, and the results indicate that the UKAN model outperforms traditional models in terms of metrics such as the intersection over union (IoU), precision, recall, and F1 score. On the C2S-MS Floods dataset and the MMFlood dataset, UKAN achieves IoUs of 87.95% and 78.31%, respectively, representing improvements of approximately 3.5 and three percentage points, respectively, over those of the traditional U-Net. Moreover, the model has significant advantages in terms of parameter efficiency and computational efficiency. These findings suggest that the UKAN model possesses greater accuracy and robustness in flood inundation area extraction tasks, which is highly important for increasing the monitoring and early warning capabilities of flood disasters.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001431025500001  
Indexed Date:  
2025-03-03  
Institution Address:  
Beijing Normal Univ, Joint Int Res Lab Catastrophe Simulat & Syst Risk, Zhuhai 519087, Peoples R China  
Institution Address:  
Beijing Normal Univ, Sch Natl Safety & Emergency Management, Zhuhai 519087, Peoples R China  
Institution:  
Institution:

### 853. Clothing attribute recognition algorithm based on improved YOLOv4-Tiny

摘要: Aiming at the problem of low accuracy of clothing attribute recognition caused by factors such as scale, occlusion and beyond the boundary, a novel clothing attribute recognition algorithm based on improved YOLOv4-Tiny is proposed in this paper. YOLOv4-Tiny is used as the basic model, firstly, the multi-scale feature extraction module Res2Net is adopted to optimize the backbone network, the receptive field size of each layer of the network is increased, and more abundant fine-grained multi-scale clothing feature information is extracted. Then, the three feature layers of the output of feature extraction network are up-sampled, and the high-level semantic features and shallow features are fused to obtain rich shallow fine-grained feature information. Finally, K-Means clustering algorithm is employed to optimize the anchor box parameters to obtain the anchor box that is more compatible with the clothing object, and to improve the integrating degree between the clothing attribute characteristics and the network. The experimental results demonstrate that the proposed method outperforms the original YOLOv4-tiny network in terms of accuracy, speed, and model parameters, and is more suitable for deployment in resource-limited embedded devices.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:000990631400001  
Indexed Date:  
2023-06-13  
Institution Address:  
Xian Polytech Univ, Sch Elect Informat, Xian 710048, Peoples R China  
Institution:

### 854. Contraband classification method for X-ray security images considering

摘要: X-ray security image contraband classification is widely used to assist in maintaining aviation and transportation security. This paper suggests an end-to-end X-ray security inspection image classification method that takes sample imbalance into account in order to address the issues of different scales of contraband in X-ray images, challenging samples, and unbalanced positive and negative samples inherent in passenger baggage security inspection. The feature fusion module is used to enhance the models ability to express picture edge and texture features while the multi-scale feature extraction network is used to capture the features of numerous sorts of illegal goods with various scales. Based on the cost-sensitive idea, the loss function is designed to solve the problem of dataset imbalance, and improve the classification accuracy of difficult samples.The experimental results of the subset constructed on the public dataset SIXray show that the proposed method improves the mean AP index by 4.5% compared with the current optimal end-to-end classification model, especially for hard-to-classify samples such as scissors, the AP index has a significant improvement effect.  
摘要:  
X光安检图像违禁品分类被广泛应用于协助维护航空和运输安全。针对X光安检图像中违禁品尺度不一、存在困难样本及旅客行李安检固有的正负样本不均衡等问题,提出一种端到端的考虑样本不平衡的X光安检图像违禁品分类方法。采用多尺度特征提取网络捕获尺度不一的多类型违禁品特征,通过特征融合模块提升模型对图像边缘和纹理特征的表达能力,基于代价敏感思想设计损失函数,解决数据集不平衡问题,并提高困难样本分类精准度。在公开数据集SIXray上构建的子集实验结果表明:所提方法相较于端到端分类模型,平均AP指标值提升了4.5%,特别是对剪刀等难分类样本,AP指标值都有显著的提升效果。  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1001-5965  
入藏号:  
CSCD:7629495  
Indexed Date:  
2024-03-15  
Institution Address:  
中国民航大学计算机科学与技术学院;;民航智慧机场理论与系统重点实验室, ;;民航智慧机场理论与系统重点实验室, ;;, 天津;;天津 300300;;300300, 中国  
Institution Address:  
School of Computer Science and Technology, Civil Aviation University of China;;Key Laboratory of Intelligent Airport Theory and System, CAAC, ;;Key Laboratory of Intelligent Airport Theory and System, CAAC, ;;, Tianjin;;Tianjin 300300;;300300  
Institution:  
中国民航大学计算机科学与技术学院;;民航智慧机场理论与系统重点实验室  
Institution:  
School of Computer Science and Technology, Civil Aviation University of China;;Key Laboratory of Intelligent Airport Theory and System, CAAC

### 855. The DeepLabV3+Algorithm Combined With the ResNeXt Network for Medical

摘要: This paper presents a semantic segmentation algorithm for medical images, leveraging the DeepLabV3+ architecture in conjunction with the ResNeXt network. The proposed algorithm takes into account the correlation between each structure of lung images and the unique characteristics of image features. Firstly, the cavity convolution algorithm is employed to enhance the receptive field of the network's feature map without augmenting the number of network parameters. Then, the extraction of dense pixel features and the expansion of the receptive field for lung images are conducted using a Densely Connected Atrous Spatial Pyramid Pooling (DenseASPP) module integrated with the ResNeXt network, which is based on multi-scale feature fusion. This ultimately leads to improved refinement of the edges in segmented lung images. The algorithm has shown excellent performance in clinical applications, providing medical professionals with more precise and accurate data to inform diagnostic and treatment strategies. Our algorithm achieved Mean Pixel Accuracy (MPA) of 0.9866, Intersection Over Union (IOU) of 0.9886 and Mean Intersection over Union (MIoU) of 0.9761, which demonstrates superiority over other state-of-the-art algorithms.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1532-0626  
eISSN:  
1532-0634  
入藏号:  
WOS:001422039500001  
Indexed Date:  
2025-02-21  
Institution Address:  
City Univ Macau, Sch Data Sci, Macau, Peoples R China  
Institution Address:  
Ningbo Univ Finance & Econ, Coll Digital Technol & Engn, Ningbo, Peoples R China  
Institution Address:  
Fuzhou Univ Int Studies & Trade, Sch Big Data, Fuzhou, Peoples R China  
Institution:  
Institution:  
Institution:  
Fuzhou Univ Int Studies & Trade

### 856. Adaptive feedback connection with a single-level feature for object

摘要: From the perspective of detector optimisation, detecting objects using only a one-level feature cannot provide good performance for a wide range of scales. Various complex feature pyramidal structures address this problem using the divide-and-conquer strategy and multi-scale feature fusion. However, this requires adding too many additional convolutional layers and fusion operations. To address the issue, a simple detection part is proposed, which includes three components, namely a one-level feature map for detection, the encoder structure with feedback connection, and a decoupled head. The redesigned encoder and decoupled head can successfully address the performance decline caused by the one-level feature-based detection. Moreover, the proposed method can accelerate the convergence of the detector and achieve a faster inference time. Based on the optimised detection part, an adaptive feedback connection with a single-level feature (AFS) is proposed for object detection. The experiments conducted on the MS COCO 2017 benchmark show that the proposed method can achieve comparable results with its multi-scale pyramid counterpart, You Only Look Once v4 (YOLOv4). In addition, AFS can help the YOLOv4 achieve 44.9 mAP at 27 frame per second and converging 82 epochs earlier under the image size of 608x608, which represents a 42.1% improvements in the convergence speed.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1751-9632  
eISSN:  
1751-9640  
入藏号:  
WOS:000821038500001  
Indexed Date:  
2022-07-10  
Institution Address:  
Chinese Acad Sci, Xian Inst Opt & Precis Mech, Xian 710119, Peoples R China  
Institution Address:  
China Aerosp Acad Syst Sci & Engn, Beijing, Peoples R China  
Institution:  
Institution:  
China Aerosp Acad Syst Sci & Engn

### 857. Hyper-Progressive Real-Time Detection Transformer (HPRT-DETR) algorithm

摘要: Aiming to address common defects such as scratches, cracks, bumps, and indentations on the surface of metal bipolar plates, this thesis proposes an algorithm called HPRT-DETR for detecting defects on metal bipolar plates. The algorithm aims to address issues such as small defects, complex backgrounds, and low detection accuracy. To enhance the performance of the algorithm, we adopt the DA to improve the AIFI module. This enhancement enables the algorithm to focus on the defective region, helping it capture more informative features. Meanwhile, we have implemented Zoom -cat scaling splicing and SSF to enhance the multi -scale feature fusion capability of the network in the CCFM module. Additionally, we have introduced the NWD metric loss to reduce sensitivity to small target locations, thereby improving detection accuracy and efficiency. Experimental validation shows that the enhanced HPRT-DETR model achieves improvements of 6.4, 1.7, and 4.7 percentage points in accuracy, recall, and average precision, respectively, compared to the original model. These results indicate that the enhanced model lays the foundation for automated production and intelligent inspection of metal bipolar plates.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
0360-3199  
eISSN:  
1879-3487  
入藏号:  
WOS:001256903000001  
Indexed Date:  
2024-07-04  
Institution Address:  
Jiangsu Univ, Mech Engn, 301 Xuefu Rd, Zhenjiang 212013, Jiangsu, Peoples R China  
Institution Address:  
Jiangsu CPTEK Serv Technol Co Ltd, 28 Shangshang Rd, Liyang 213300, Jiangsu, Peoples R China  
Institution:  
Institution:  
Jiangsu CPTEK Serv Technol Co Ltd

### 858. EGBNet: a real-time edge-guided bilateral network for nighttime semantic

摘要: Due to poor illumination and low contrast, semantic segmentation of nighttime images faces major challenges. Various segmentation models with a large number of parameters are proposed to improve the performance but lead to an inability to process in real time. To tackle these problems, we propose a real-time edge-guided bilateral network (EGBNet) for nighttime semantic segmentation. Considering the blurred details and low contrast of nighttime images, we propose a lightweight multi-dilation dense aggregation module and introduce an efficient edge head to improve the ability to distinguish target features from the nighttime background. Moreover, a self-adaptive feature fusion module is proposed for the bilateral segmentation network to enhance the feature representation and generalization ability by fully using multi-scale feature maps. To capture more useful information from limited nighttime images, we further use the knowledge distillation strategy to improve the segmentation performance. Extensive experiments on ACDC and BDD datasets demonstrate the effectiveness of our EGBNet by achieving a satisfactory trade-off between segmentation accuracy and inference speed. Specifically, EGBNet achieves 55.56% mIoU on the ACDC test set with 9.4 M parameters and 60FPS speed for a 1080 x 1920 input image on a single NVIDIA 2080Ti.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:000956006300001  
Indexed Date:  
2023-04-13  
Institution Address:  
Tianjin Univ, Sch Elect & Informat Engn, Tianjin, Peoples R China  
Institution:

### 859. Multi-Scale Feature Attention Fusion for Image Splicing Forgery

摘要: Image splicing is a widely occurrence image tampering technology. With the rapid development of digital image processing technology, detecting image splicing forgery has become significantly challenging. Although various methods have been devised to identify such tampered images, existing approaches have not achieved optimal performance due to limitations in effectively leveraging feature maps of different scales. To address this issue, we propose a novel method for image splicing forgery detection called multi-scale feature attention fusion network (MFAF-Net). We propose a multi-scale atrous feature attention (MAFA) module designed to capture rich contextual features for multi-scale high-level feature fusion. Additionally, we present the multi-branch attention mechanism (MBAM) module to fuse contextual information from various branches for low-level features. This integration enhances the capability of low-level features to produce more refined pixel-level attention. We employ the weighted binary cross-entropy loss and dice loss in the MFAF-Net to overcome the imbalance between positive and negative samples. Extensive experiments demonstrate that the proposed MFAF-Net outperforms state-of-the-art methods. Robustness experiments also show our model exhibits image splicing forgery detection robustness under common attacks. CCS Concepts: center dot Computing methodologies -> Image processing; center dot Security and privacy -> Social aspects of security and privacy;  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1551-6857  
eISSN:  
1551-6865  
入藏号:  
WOS:001415998900001  
Indexed Date:  
2025-02-15  
Institution Address:  
Harbin Inst Technol, Sch Comp Sci & Technol, Shenzhen, Peoples R China  
Institution Address:  
City Univ Hong Kong, Dept Comp Sci, Hong Kong, Peoples R China  
Institution:  
Institution:

### 860. Tripartite Feature Enhanced Pyramid Network for Dense Prediction

摘要: Learning pyramidal feature representations is important for many dense prediction tasks (e.g., object detection, semantic segmentation) that demand multi-scale visual understanding. Feature Pyramid Network (FPN) is a well-known architecture for multi-scale feature learning, however, intrinsic weaknesses in feature extraction and fusion impede the production of informative features. This work addresses the weaknesses of FPN through a novel tripartite feature enhanced pyramid network (TFPN), with three distinct and effective designs. First, we develop a feature reference module with lateral connections to adaptively extract bottom-up features with richer details for feature pyramid construction. Second, we design a feature calibration module between adjacent layers that calibrates the upsampled features to be spatially aligned, allowing for feature fusion with accurate correspondences. Third, we introduce a feature feedback module in FPN, which creates a communication channel from the feature pyramid back to the bottom-up backbone and doubles the encoding capacity, enabling the entire architecture to generate incrementally more powerful representations. The TFPN is extensively evaluated over four popular dense prediction tasks, i.e., object detection, instance segmentation, panoptic segmentation, and semantic segmentation. The results demonstrate that TFPN consistently and significantly outperforms the vanilla FPN. Our code is available at https://github.com/jamesliang819.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
35  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
36  
ISSN:  
1057-7149  
eISSN:  
1941-0042  
入藏号:  
WOS:001025042000001  
PubMed ID:  
37155388  
Indexed Date:  
2023-07-24  
Institution Address:  
Rochester Inst Technol, Dept Comp Engn, Rochester, NY 14623 USA  
Institution Address:  
Univ Rochester, Dept ECE & CS, Rochester, NY 14642 USA  
Institution Address:  
Swiss Fed Inst Technol, Comp Vis Lab, CH-8092 Zurich, Switzerland  
Institution:  
Institution:  
Institution:

### 861. LBARNet: Lightweight bilateral asymmetric residual network for real-time

摘要: Real-time semantic segmentation, as a key technique for scene understanding, has been an important research topic in the field of computer vision in recent years. However, existing models are unable to achieve good segmentation accuracy on mobile devices due to their huge computational overhead, which makes it difficult to meet actual industrial requirements. To address the problems faced by current semantic segmentation tasks, this paper proposes a lightweight bilateral asymmetric residual network (LBARNet) for real-time semantic segmentation. First, we propose the bilateral asymmetric residual (BAR) module. This module learns multi-scale feature representations with strong semantic information at different stages of the semantic information extraction branch, thus improving pixel classification performance. Secondly, the spatial information extraction (SIE) module is constructed in the spatial detail extraction branch to capture multi-level local features of the shallow network to compensate for the lost geometric information in the downsampling stage. At the same time, we design the attention mechanism perception (AMP) module in the jump connection part to enhance the contextual representation. Finally, we design the dual branch feature fusion (DBF) module to exploit the correspondence between higher-order features and lower-order features to fuse spatial and semantic information appropriately. The experimental results show that LBARNet, without any pre-training and pre-processing and using only 0.6M parameters, achieves 70.8% mloU and 67.2% mloU on the Cityscapes dataset and Camvid dataset, respectively. LBARNet maintain a high segmentation accuracy while using a smaller number of parameters compared to most existing state-of-the-art models.& COPY; 2023 Elsevier Ltd. All rights reserved.  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
10  
ISSN:  
0097-8493  
eISSN:  
1873-7684  
入藏号:  
WOS:001061926900001  
Indexed Date:  
2023-09-25  
Institution Address:  
Chongqing Univ Posts & Telecommun, Sch Commun & Informat Engn, Chongqing 400065, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Chongqing Key Lab Signal & Informat Proc, Chongqing 400065, Peoples R China  
Institution:  
Institution:

### 862. FDLR-Net: A feature decoupling and localization refinement network for

摘要: Object detection in remote sensing images is a critical task in computer vision. Often times in remote sensing images, objects are highly variable in scale and have arbitrary orientation, which renders spatial alignment between anchor boxes and objects challenging in the object detection task. In this paper, a feature decoupling and localization refinement network is suggested as a solution to this issue. Specifically, a bidirectional feature fusion module (BFFM) is devised to construct a multi-scale feature pyramid for detecting objects at different scales. A feature decoupling module (FDM) is devised which utilizes the fusion of spatial attention and channel attention, as well as different attention functions to generate features specifically tuned for regression and classification, that are used to guide more accurate localization and classification. Further, a localization refinement module (LRM) is designed to automatically optimize the anchor box parameters to achieve spatial alignment of the anchor box and the object regression feature. In this way, the FDM and LRM are cascaded to achieve more accurate localization. Experimental results on two open access datasets, DOTA and HRSC2016, show that the performance of FDLR-Net is state-of-the-art, with mAP reaching 73.08% and 89.4%, respectively.  
Web of Science 核心合集中的 "被引频次":  
24  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
25  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:000984539200001  
Indexed Date:  
2023-05-29  
Institution Address:  
Wuhan Univ, Sch Elect Informat, Wuhan 430072, Peoples R China  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying, Mapping & Remote Sensing, Wuhan 430072, Peoples R China  
Institution Address:  
Univ Calif Santa Barbara, Dept Comp Sci, Santa Barbara, CA 93106 USA  
Institution:  
Institution:  
Institution:

### 863. Ship target detection method for synthetic aperture radar images based

摘要: Aiming at the problem that target detection in synthetic aperture radar(SAR)images is easily affected by noise and background interference,and the performance of ship target detection is degraded under multi-scale conditions,an improved YOLOv5algorithm is proposed on the basis of considering the network scale and detection accuracy.In this algorithm,coordinate attention mechanism is used to suppress noise and interference to improve the feature extraction ability of the network while ensuring its lightweight advantage. The bi-directional feature pyramid is integrated to achieve multi-scale feature fusion.A new prediction box loss function is designed to improve the detection accuracy and accelerate the convergence of the algorithm.Thus, the ship target can be recognized quickly and accurately in SAR images.Experimental verification shows that the mean average presicion(mAP)of the proposed algorithm on SSDD dataset reaches 96.7%,which is 1.9% higher than that of YOLOv5s.The convergence speed is faster during training,and the network is lightweight, which has a good prospect in practical application.  
摘要:  
针对合成孔径雷达图像目标检测易受噪声和背景干扰影响,以及多尺度条件下检测性能下降的问题,在兼顾网络规模和检测精度的基础上,提出了一种改进的合成孔径雷达舰船目标检测算法。使用坐标注意力机制,在确保轻量化的同时抑制了噪声与干扰,以提高网络的特征提取能力;融入加权双向特征金字塔结构以实现多尺度特征融合,设计了一种新的预测框损失函数以改善检测精度,同时加快算法收敛,从而实现了对合成孔径雷达图像舰船目标的快速准确识别。实验验证表明,所提算法在合成孔径雷达舰船检测数据集(synthetic aperture radar ship detection dataset,SSDD)上的平均精度均值达到96.7%,相比于YOLOv5s提高1.9%,训练时收敛速度更快,且保持了网络轻量化的特点,在实际应用中具有良好前景。  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1001-506X  
入藏号:  
CSCD:7618589  
Indexed Date:  
2024-03-02  
Institution Address:  
火箭军工程大学作战保障学院;;国防科技大学信息与通信学院, ;;, 西安;;武汉, 陕西;;湖北 710025;;430010, 中国  
Institution Address:  
火箭军工程大学作战保障学院, 西安, 陕西 710025, 中国  
Institution Address:  
Combat Support College,Rocket Force University of Engineering;;College of Information and Communication,National University of Defense Technology, ;;, Xian;;Wuhan, ;; 710025;;430010  
Institution Address:  
Combat Support College,Rocket Force University of Engineering, Xian, 710025  
Institution:  
火箭军工程大学作战保障学院;;国防科技大学信息与通信学院  
Institution:  
火箭军工程大学作战保障学院  
Institution:  
Combat Support College,Rocket Force University of Engineering;;College of Information and Communication,National University of Defense Technology  
Institution:  
Combat Support College,Rocket Force University of Engineering

### 864. Adaptive Classification Network for Similar Features Between Classes in

摘要: Addressing the issue of inter-class similarity is a challenging task in the research of autonomous driving scene classification, which primarily focuses on learning the distinctive features of targets in real-world complex traffic scenarios with high similarity, and constructing the overall correlation between features for scene classification. To this end, a multi-scale adaptive feature selection network for autonomous driving scene classification is proposed. Initially, a dual multi-scale feature extraction module is utilized for preliminary processing to extract inter-class similar features at different scales. Subsequently, a feature differentiation screening module is designed to complete the screening of scene-similar features, enabling the network to focus more on the typical and easily distinguishable features of different scene categories. Then, the feature screening results and multi-scale feature maps are transferred to the feature fusion classification module for scene classification, and the correlation between scene features is captured. Finally, an adaptive learning algorithm dynamically adjusts the training parameters through the output results, accelerating the network's convergence speed and improving accuracy. The proposed method is compared with existing network methods on three datasets: BDD100k, BDD100k + and self-made dataset. Compared with the Top2 networks, it leads in accuracy by 3.29%, 5.59% and 12.65% (relatively), respectively. Experimental results demonstrate the effectiveness of the proposed method and its strong generalization capability. The scene classification method presented in this paper aims to learn the typical and easily distinguishable features and their correlations under different complex scene categories, reducing the impact of inter-class similarity among multiple targets, thereby making the scene classification results in real-world traffic scenario datasets more accurate.  
摘要:  
解决类间相似度问题是自动驾驶场景分类研究中一项充满挑战的任务,主要研究在相似度较高的真实复杂交通场景中,利用网络学习目标特征的差异性,并构建特征之间整体关联性进行场景分类。提出一种多尺度自适应特征筛选的自动驾驶场景分类网络。采用双重多尺度特征提取模块预处理,初步提取不同尺度下的类间相似特征;设计了特征分化筛选模块完成场景相似特征筛选,使网络更关注不同场景类别的典型易区分特征;将特征筛选结果和多尺度特征图共同传递至特征融合分类模块进行场景分类,捕捉场景特征之间的关联性;由自适应学习算法通过输出结果动态调整训练参数,加快网络收敛速度并提升精度。所提方法在三种数据集BDD100k、BDD100k+和自制数据集上与现有网络方法进行比较,相较Top2网络在精度上分别领先了3.29%、5.59%、12.65%(相对),实验结果表明了所提方法的有效性,并展现了很好的泛化能力。提出的场景分类方法旨在学习不同复杂场景类别下的典型易区分的特征及其关联性,降低多目标类间相似的影响,使得在真实交通场景数据集中场景分类结果更加准确。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1673-9418  
入藏号:  
CSCD:7836794  
Indexed Date:  
2025-01-11  
Institution Address:  
辽宁工程技术大学软件学院;;优策(江苏)安全科技有限公司OpenSafe实验室, ;;, 葫芦岛;;苏州, 辽宁;;江苏 125105;;215100, 中国  
Institution Address:  
优策(江苏)安全科技有限公司OpenSafe实验室;;清华大学苏州汽车研究院(相城), ;;, 苏州;;苏州, 江苏;;江苏 215100;;215100, 中国  
Institution Address:  
辽宁工程技术大学软件学院, 葫芦岛, 辽宁 125105, 中国  
Institution Address:  
College of Software,Liaoning Technical University;;OpenSafe Lab,Utcer (Jiangsu) Safety Technology Co.,Ltd., ;;, Huludao;;Suzhou, Liaoning;;Jiangsu 125105;;215100  
Institution Address:  
OpenSafe Lab,Utcer (Jiangsu) Safety Technology Co.,Ltd.;;Suzhou Automobile Research Institute,Tsinghua University, ;;, Suzhou;;Suzhou, Jiangsu;;Jiangsu 215100;;215100  
Institution Address:  
College of Software,Liaoning Technical University, Huludao, Liaoning 125105, China  
Institution:  
辽宁工程技术大学软件学院;;优策(江苏)安全科技有限公司OpenSafe实验室  
Institution:  
优策(江苏)安全科技有限公司OpenSafe实验室;;清华大学苏州汽车研究院(相城)  
Institution:  
辽宁工程技术大学软件学院  
Institution:  
College of Software,Liaoning Technical University;;OpenSafe Lab,Utcer (Jiangsu) Safety Technology Co.,Ltd.  
Institution:  
OpenSafe Lab,Utcer (Jiangsu) Safety Technology Co.,Ltd.;;Suzhou Automobile Research Institute,Tsinghua University  
Institution:  
College of Software,Liaoning Technical University

### 865. Aggregating discriminative embedding by triple-domain feature joint

摘要: Each axis of the speech, including time-domain, frequency-domain, and spectral-domain data, that's represents different physical meanings and different dimension information. Time-domain focus on physical signal versus time, the frequency-domain focus on the amount of signal in a given frequency band, and the spectral-domain focus on global power with speech, so only using the spectrogram to represent the whole information of speech to do speaker recognition will lose a lot of details information with the other dimensions. To tackle this limitation, we propose a triple-domain feature joint learning to enhance discriminative embedding from more dimensions for text-independent speaker verification. To further aggregate discriminative embedding, each domain uses a novel bidirectional sampling multi-scale feature aggregation network based on Fisher feature fusion to project spectrum features to more discriminative embeddings, termed TribiNet. Extensive experiments are conducted on a text-independent speaker verification dataset generated from the VoxCeleb corpus. The re-sults demonstrate that the proposed method outperforms the state-of-the-art deep embedding architectures by at least 12%-58% in the test set. The ablation experiments further illustrate that our proposed approaches can achieve substantial improvement over prior methods.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:000991663000001  
Indexed Date:  
2023-06-07  
Institution Address:  
Wuhan Univ Technol, Sch Comp & Artificial Intelligence, 122 Luoshi Rd, Wuhan 430070, Peoples R China  
Institution Address:  
Wuhan Univ Technol, Sanya Sci & Educ Innovat Pk, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Technol, Sch Comp & Artificial Intelligence, Sanya Sci & Educ Innovat Pk, Wuhan, Peoples R China  
Institution:  
Institution:  
Institution:

### 866. New technology of seismic random noise suppression based on DMFF-Net

摘要: Seismic exploration is one of the most widely used geophysical prospecting methods in oil-gas and mineral resources development. Due to the limitations of acquisition conditions, the seismic records are usually contaminated with a large amount of random noise, resulting in a low Signal-to-Noise Ratio (SNR). It seriously affects the identification accuracy of the effective signals, thereby bringing challenges to subsequent inversion and interpretation procedures. In addition, the random noise usually has complex characteristics, such as non-stationary, non-Gaussian and spectral aliasing. The denoising performance for the conventional methods may degrade when confronted with such complex interferences. To achieve the complex noise attenuation, a novel double-layer multi-scale feature fusion denoising network (DMFF-Net) is proposed in this paper. In general, the proposed network has a multi-scale network structure. It utilizes the multi-branch modules to extract the potential features existing in different scales and branches so as to improve the learning ability of the network for complex features of the analyzed seismic data. Meanwhile, we also employ skip connections to fuse the shallow and deep features; then, improve the recover ability of the weak signals. The synthetic and field data processing results indicate that DMFF-Net can suppress the random noise effectively and restore the desired signals accurately. Moreover, it also can significantly improve the SNR. Compared with conventional denoising methods, DMFF-Net has advantages in signal amplitude retention and weak signal recovery.  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0001-5733  
入藏号:  
WOS:000893200500022  
Indexed Date:  
2022-12-28  
Institution Address:  
Minist Educ, Key Lab Modern Power Syst Simulat & Control & Ren, Jilin 132012, Jilin, Peoples R China  
Institution Address:  
Northeast Elect Power Univ, Dept Commun Engn, Jilin 132012, Jilin, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Sch Earth Sci & Engn, Guangzhou 510275, Guangdong, Peoples R China  
Institution Address:  
Southern Marine Sci & Engn Guangdong Lab, Zhuhai 519000, Guangdong, Peoples R China  
Institution Address:  
Jilin Univ, Coll Instrumentat & Elect Engn, Changchun 130026, Peoples R China  
Institution Address:  
Jilin Univ, Coll Commun Engn, Changchun 130012, Peoples R China  
Institution Address:  
Jilin Univ, Coll Geoexplorat Sci & Technol, Changchun 130026, Peoples R China  
Institution:  
Minist Educ  
Institution:  
Institution:  
Institution:  
Southern Marine Sci & Engn Guangdong Lab  
Institution:  
Institution:  
Institution:

### 867. Multi-scale feature extraction and fusion with attention interaction for

摘要: RGB-T single-object tracking aims to track objects utilizing both RGB images and thermal infrared(TIR) images. Though the siamese-based RGB-T tracker shows its advantage in tracking speed, its accuracy still cannot be compared with other state-of-the-art trackers (e.g., MDNet). In this study, we revisit the existing siamese-based RGB-T tracker and find that such fall behind comes from insufficient feature fusion between RGB image and TIR image, as well as incomplete interactions between template frame and search frame. Inspired by this, we propose a multi-scale feature extraction and fusion network with Temporal-Spatial Memory (MFATrack). Instead of fusing RGB image and TIR image with the single-scale feature map or only high-level features from the multi-scale feature map, MFATrack proposes a new fusion strategy by fusing features from all scales, which can capture contextual information in shallow layers and details in the deep layer. To learn the feature better for tracking tasks, MFATrack fuses the features via several consecutive frames. In addition, we also propose a self-attention interaction module specifically designed for the search frame, highlighting the features in the search frame that are relevant to the target and thus facilitating rapid convergence for target localization. Experimental results demonstrate the proposed MFATrack is not only fast, but also can obtain better tracking accuracy compared with other competing methods including MDNet-based methods and other siamese-based trackers.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:001304116300001  
Indexed Date:  
2024-09-07  
Institution Address:  
Northwestern Polytech Univ, Sch Comp Sci, Xian 710072, Peoples R China  
Institution:

### 868. Multi-scale fusion and efficient feature extraction for enhanced sonar

摘要: Sonar imaging is an underwater detection technology that relies on the transmission and reception of acoustic pulse waves. This technique plays a crucial role in various domains including underwater archaeology, energy exploration, and oceanographic surveying. A primary challenge in sonar imaging is the low signalto-noise ratio and significant noise interference, issues that are influenced by the constraints of equipment performance and the underwater environment. Traditional object detection techniques have been inadequate in effectively extracting deep features from sonar images possessing targets with complex structures and have also demonstrated shortcomings in the fusion processing of target features at multiple scales, thereby affecting the robustness and accuracy of object detection. To improve the performance of sonar image object detection, we propose an advanced detection framework that integrates efficient feature extraction with multi-scale feature fusion. We employ EfficientNet as the backbone network. EfficientNet exhibits excellent feature extraction capabilities through comprehensive adjustments of depth, width, and resolution. We introduce a dual-channel attention module that blends Squeeze-and-Excitation (SE) and Efficient Channel Attention (ECA) mechanisms to amplify the expression of crucial feature channels and suppress the lesser ones. Additionally, we utilize a modified bidirectional feature pyramid network (BiFPN) to strengthen the integration of features across different layers. Employing these methods, we amalgamate the features into a shared weights classification network and bounding box prediction network for accurate target class discernment and localization. The experimental outcomes provide evidence of the notable superlative nature of the proposed framework in sonar image object detection, effectively ameliorating detection performance amidst noise interference.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001291502800001  
Indexed Date:  
2024-08-18  
Institution Address:  
Hohai Univ, Coll Artificial Intelligence & Automat, Changzhou 213200, Jiangsu, Peoples R China  
Institution Address:  
Hubei Technol Innovat Ctr Smart Hydropower, Wuhan 430000, Peoples R China  
Institution Address:  
Hohai Univ, Coll Informat Sci & Engn, Changzhou 213200, Jiangsu, Peoples R China  
Institution:  
Institution:  
Hubei Technol Innovat Ctr Smart Hydropower  
Institution:

### 869. YOLO-CEA: a real-time industrial defect detection method based on

摘要: This paper proposes a real-time industrial defect detection method based on context enhancement and attention to address the problem that current general-purpose target detectors can hardly achieve high detection accuracy and fast detection speed simultaneously. First, a modified MonileNetV3 is used as the backbone network to reduce the number of parameters and improve the model detection speed. A lightweight TRANS module is proposed at the end of the backbone network to combine more layers of features provided by global contextual information for complex background small target detection. Secondly, a cross-layer multi-scale feature fusion network is designed to fully fuse the fine-grained and semantic feature information extracted by the backbone and enhance the spatial location information between neighboring feature layers. Finally, a cascaded Two-channel Efficient Space attention module is used to fully extract texture and semantic features from the defective regions, allowing the model to focus more on the wrong locations and improve the feature representation capability of the network. The NEU-DET steel and PCB datasets are used to test the effectiveness of the proposed model. The experimental results show that compared to the original YOLOv5s algorithm, the mAP metrics are improved by 5.9% and 0.6%, F1 is improved by 4.82% and 0.93%, respectively, and the parameters are reduced by 33.77 M, enabling fast detection of industrial surface defects and meeting the needs of the entire industry.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1386-7857  
eISSN:  
1573-7543  
入藏号:  
WOS:001017749800001  
Indexed Date:  
2023-07-08  
Institution Address:  
Qilu Univ Technol, Shandong Acad Sci, Shandong Comp Sci Ctr, Natl Supercomp Ctr Jinan, Jinan 250353, Peoples R China  
Institution:

### 870. Efficient Livestock Detection in Grazing Areas Based on Enhanced

摘要: Realizing big data to manage livestock requires real-time monitoring of livestock,but real-time monitoring of livestock is easily interfered by large changes in target size,lighting,environmental factors,etc.,so it is difficult to detection,and existing livestock detection algorithms have the problem of poor robustness.An object detection network called E-YOLOv4-tiny is proposed based on enhanced YOLOv4-tiny,which adopts a pyramid network with multi-scale feature fusion,taking into account shallow local detail features and deep semantic information to solve the problem of livestock size fluctuation in pastoral areas. The number of backbone network parameters is reduced by improving the residual structure to accommodate embedded platform requirements.A new composite clustering algorithm is introduced to design anchor frames to improve the accuracy of the algorithm under the premise of ensuring portability.Finally,according to the characteristics of a pastoral environment,a new Compound Muti-channel Attention(CMA) mechanism is proposed to improve the poor accuracy of the target detection network and enhance the robustness of the algorithm.Experimental results show that the mean Average Precision(mAP) of the E-YOLOv4-tiny algorithm is 0.878 9,and the frame rate is 32 frame/s, and it's mAP is 9.32% higher than that of the traditional YOLOv4-tiny algorithm while maintaining almost the same detection rate.  
摘要:  
实现大数据管理牲畜需要实时监测牲畜,但对牲畜进行实时监测容易受到目标尺寸变化大、光照、环境因素等干扰,因此检测难度大,现有牲畜检测算法存在鲁棒性差等问题。提出一种基于增强型YOLOv4-tiny的目标检测算法(E-YOLOv4-tiny),采用多尺度特征融合的金字塔网络,兼顾浅层局部细节特征与深层语义信息,解决牧区牲畜尺寸波动问题。通过改进残差结构,减少主干网络参数量,以适应嵌入式平台需求。引入一种新的复合聚类算法设计锚框,在保证可移植性的前提下提高算法精度。针对牧区环境特点,提出一种新的复合多通道注意力机制,改善目标检测网络精度差的问题,增强算法鲁棒性。实验结果表明,E-YOLOv4-tiny算法的平均精度均值(mAP)为0.878 9,帧率为32帧/s,相较于传统YOLOv4-tiny算法,在保持几乎相同的检测速率条件下,mAP提升了9.32%。  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1000-3428  
入藏号:  
CSCD:7509558  
Indexed Date:  
2023-10-28  
Institution Address:  
内蒙古自治区机电控制重点实验室;;内蒙古工业大学电力学院, 内蒙古自治区机电控制重点实验室;;, 呼和浩特;;呼和浩特, ;; 010051;;010051  
Institution Address:  
内蒙古自治区机电控制重点实验室, 内蒙古自治区机电控制重点实验室, 呼和浩特, 内蒙古 010051, 中国  
Institution Address:  
内蒙古自治区机电控制重点实验室;;水利部牧区水利科学研究所, 内蒙古自治区机电控制重点实验室;;, 呼和浩特;;呼和浩特, ;; 010051;;010051  
Institution Address:  
内蒙古北方龙源风力发电有限责任公司, 呼和浩特, 内蒙古 010050, 中国  
Institution Address:  
Inner Mongolia Key Laboratory of Electrical and Mechanical Control;;Institute of Electric Power,Inner Mongolia University of Technology, Inner Mongolia Key Laboratory of Electrical and Mechanical Control;;, Hohhot;;Hohhot, ;; 010051;;010051  
Institution Address:  
Inner Mongolia Key Laboratory of Electrical and Mechanical Control, Inner Mongolia Key Laboratory of Electrical and Mechanical Control, Hohhot, Inner Mongolia 010051, China  
Institution Address:  
Inner Mongolia Key Laboratory of Electrical and Mechanical Control;;Pastoral Water Conservancy Science Research Institute of the Ministry of Water Resources, Inner Mongolia Key Laboratory of Electrical and Mechanical Control;;, Hohhot;;Hohhot, ;; 010051;;010051  
Institution Address:  
Inner Mongolia North Longyuan Wind Power Generation Co.,Ltd., Hohhot, Inner Mongolia 010050, China  
Institution:  
内蒙古自治区机电控制重点实验室;;内蒙古工业大学电力学院  
Institution:  
内蒙古自治区机电控制重点实验室  
Institution:  
内蒙古自治区机电控制重点实验室;;水利部牧区水利科学研究所  
Institution:  
内蒙古北方龙源风力发电有限责任公司  
Institution:  
Inner Mongolia Key Laboratory of Electrical and Mechanical Control;;Institute of Electric Power,Inner Mongolia University of Technology  
Institution:  
Inner Mongolia Key Laboratory of Electrical and Mechanical Control  
Institution:  
Inner Mongolia Key Laboratory of Electrical and Mechanical Control;;Pastoral Water Conservancy Science Research Institute of the Ministry of Water Resources  
Institution:  
Inner Mongolia North Longyuan Wind Power Generation Co.,Ltd.

### 871. Improved YOLOv8 Algorithm for Industrial Surface Defect Detection

摘要: Aiming at the problems of low contrast of industrial defects and high false detection rate and leakage rate caused by the surrounding interference information,it proposes an industrial surface defect detection algorithm EML-YOLO based on the improvement of YOLOv8.By designing a high-efficiency large convolution module ELK,the model's feature extraction capability can be improved by providing a multi-scale feature representation while retaining the spatial information;by proposing a parallel multi-branch feature fusion module MCM,which enables the model to acquire rich feature information and global context information;and reducing the number of parameters and computation of the model by feature compression and streamlining in the Neck module,which makes the model more applicable to industrial scenarios with limited resources.Two industrial surface defect datasets,GC10-DET and DeepPCB,are used to validate the effectiveness of the improved EML-YOLO algorithm.The experimental results show that on the GC10-DET dataset and DeepPCB dataset,the detection accuracy is improved by 4.3 percentage points and 2.9 percentage points,respectively,and the number of parametric quantities is only 2.7\*10~6.The proposed algorithm can be better applied to industrial defect detection scenarios.  
摘要:  
针对工业缺陷对比度低、周围干扰信息多导致的误检率和漏检率高的问题,提出一种基于改进YOLOv8的工业表面缺陷检测算法EML-YOLO。通过设计一种高效大卷积模块(efficient large kernel,ELK),在保留空间信息的同时提供多尺度的特征表示,从而提高模型的特征提取能力;提出多支路并行的特征融合模块(multi-scale context module,MCM),使得模型能够获取丰富的特征信息和全局上下文信息;在Neck模块中通过特征压缩和精简来减少模型的参数量和计算量,让模型更适用于资源有限的工业场景。采用GC10-DET和DeepPCB两个工业表面缺陷数据集来验证改进的EML-YOLO算法的有效性。实验结果表明,在GC10-DET数据集和DeepPCB数据集上,检测准确率上分别提高了4.3个百分点和2.9个百分点,参数量仅2.7\*10~6。所提算法可以较好地应用于工业缺陷检测场景。  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1002-8331  
入藏号:  
CSCD:7794594  
Indexed Date:  
2024-11-01  
Institution Address:  
河北科技大学信息科学与工程学院, 石家庄, 河北 050018, 中国  
Institution Address:  
School of Information Science and Engineering,Hebei University of Science and Technology, Shijiazhuang, Hebei 050018, China  
Institution:  
河北科技大学信息科学与工程学院  
Institution:  
School of Information Science and Engineering,Hebei University of Science and Technology

### 872. Image Feature Matching Method of High-Speed Railway Catenary with

摘要: Aiming at the problem that the traditional multi-scale feature matching algorithm is difficult to maintain the image local accuracy and edge details in the process of high-speed railway catenary image matching detection, an improved accelerated nonlinear diffusion (AKAZE) algorithm for high-speed railway catenary image feature matching is proposed. Firstly, the method of edge feature and local binary pattern texture feature fusion is used to overcome the shortage of feature points in traditional catenary image. Then, the improved AKAZE algorithm is used to extract the features of catenary image, and the binary robust independent elementary feature (BRIEF) descriptor is proposed to describe the feature points. Next, the false matching points are eliminated by fast similar neighborhood search and random sampling consistent algorithm. Finally, the image difference method is used to realize the matching detection of catenary image. Experimental results show that, compared with the AKAZE feature matching algorithm, the average matching accuracy of the proposed algorithm is improved by 22. 16%, and the operation efficiency of the algorithm is also greatly improved.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1006-4125  
入藏号:  
WOS:000823175200013  
Indexed Date:  
2022-07-21  
Institution Address:  
Lanzhou Jiaotong Univ, Sch Elect & Informat Engn, Lanzhou 730070, Gansu, Peoples R China  
Institution Address:  
Gansu Prov Engn Res Ctr Artificial Intelligence &, Lanzhou 730070, Gansu, Peoples R China  
Institution:  
Institution:  
Gansu Prov Engn Res Ctr Artificial Intelligence &

### 873. Global attention module and cascade fusion network for steel surface

摘要: Steel surface defect detection plays a pivotal role in contemporary society, ensuring quality and safety in construction and manufacturing, reducing production costs, improving efficiency, and driving technological innovation. However, this task encounters challenges, including addressing unstructured features, multi-scale issues, and a scarcity of available data. To overcome these challenges, this paper proposes a global attention module and cascade fusion network for steel surface defect detection, called GC-Net. In this network, the global attention module is proposed to enhance the capability of the model to handle unstructured defects. Subsequently, a cascade fusion network is designed for multi-scale feature fusion, thereby improving detection accuracy for defects of varying scales. Following this, soft non-maximum suppression is applied in the post- processing stage to eliminate redundant detection boxes, further enhancing the detection performance of the network. Finally, a series of data augmentation techniques, including oversampling and small object augmentation, are employed in the experimental sessions to mitigate the issue of data scarcity. The experimental results on two datasets for steel surface defect detection demonstrate that the proposed method outperforms state-of-the-art methods in terms of mAP50 metric (NEU-DET: 0.771, GC10-DET: 0.635).  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:001312659700001  
Indexed Date:  
2024-09-23  
Institution Address:  
Univ Sci & Technol Liaoning, Sch Elect & Informat Engn, Anshan 114051, Peoples R China  
Institution:

### 874. A novel percussion-based approach for pipeline leakage detection with

摘要: Pipelines are susceptible to oil and gas leaks during long-distance transportation due to factors such as damage from external forces and aging. However, existing pipeline leakage detection technologies that rely on physical inspections or sensors installed on pipelines are time-consuming and costly. In this paper, a percussion approach based on improved MobileNetV2 is proposed for pipeline leakage detection. Firstly, the influence of pipe leakage size on vibration characteristics was investigated by theoretical analysis and numerical simulation. Subsequently, experiments were conducted to assess the validity of the proposed method. The sounds produced by hammering the pipe under different damage conditions were recorded using a smartphone. The improved MobileNetV2 model was then used for classifying Mel spectrogram and Mel frequency cepstrum coefficient (MFCC) features extracted from the recorded sound signals. This model incorporates a multi-scale feature fusion module, which allows it to capture features at different scales and enhances its ability to differentiate between damage conditions. Experimental results show that using Mel spectrogram as input for the improved MobileNetV2 achieves a higher accuracy compared to using MFCC, with 100% accuracy for identifying leakage damage and 99.87% for classifying leak size. Compared to other methods, the improved MobileNetV2 exhibits superior classification performance while maintaining the lightweight characteristics of the original MobileNetV2. In conclusion, the improved model demonstrates significant enhancements in classification performance and operational efficiency, making it a promising approach for processing percussive signals.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001240108100001  
Indexed Date:  
2024-06-11  
Institution Address:  
Yangtze Univ, Sch Urban Construct, Jingzhou 434023, Peoples R China  
Institution Address:  
Tongji Univ, Coll Civil Engn, Shanghai 200092, Peoples R China  
Institution:  
Institution:

### 875. Multiscale Difference Feature-Fusion Network for Change Detection With

摘要: Land-cover change detection with hyperspectral remote sensing images (HyperCD) has become attractive in the applications of remote sensing images. Many existing studies have indicated that attention mechanisms play an important role in HyperCD. However, methods based on attention enhancement for HyperCD require further improvement. In this letter, we propose a novel multiscale difference feature-fusion network (MDFN) to improve the detection performance of HyperCD. First, a submodule named multiattention feature enhancement (MAFE) module was designed and embedded on each scale in the backbone of the proposed MDFN to capture subtle changes. Second, with the motivation of exploring the feature connection of a target on different scales, the attention feature maps from each scale were fused via a proposed novel cross-scale residual fusion module (CS-RFM). Finally, a softmax function was adopted to generate a binary change detection map based on the fused features. Experimental results based on comparison with five existing related works indicated that the proposed MDFN not only has some advantages in improving change detection performance with real hyperspectral remote sensing images (HRSIs) but also exhibits superiority in the requirement of training samples that are preferred in practical applications. For instance, using only 5% of the training samples, the average accuracy (AA) on the Farmland dataset is improved by 0.63%. The code will be available at https://github.com/ImgSciGroup/2024-MDFN.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001420491600007  
Indexed Date:  
2025-02-20  
Institution Address:  
Xian Univ Technol, Sch Comp Sci & Engn, Xian 710048, Peoples R China  
Institution:

### 876. CMPFFNet: Cross-Modal and Progressive Feature Fusion Network for RGB-D

摘要: Depth information can contribute to the semantic segmentation of scenes from red-green-blue (RGB) images. Therefore, the amount of information that can be obtained from RGB and RGB-depth (RGB-D) images is significantly greater for this task. However, RGB and RGB-D modalities are different in terms of object representation. Features that are extracted from these modalities and fused effectively are key to scene semantic segmentation. In addition, complete segmentation requires the fusion of multiscale features to unify global information. However, existing approaches primarily use multiscale features for sequential integration. This study introduces a cross-modal and progressive feature fusion network (CMPFFNet) for semantic segmentation of indoor scenes in RGB-D images. First, a multimodal adaptive alignment fusion (MAAF) module based on an attention mechanism is introduced. This module aligns the two modal channels by additive attention and then computes the spatial similarity between the two modalities based on the dot product to incorporate the complementary information of the depth modality into the RGB modality. In addition, a reverse attention augmentation (RAA) module is introduced to augment the more abstract high-level features for two adjacent multilevel features using the concrete semantic information of the lower-level features in them. After augmenting the extracted multilevel features, a multilevel feature progressive fusion (MFPF) module is deployed; this module sequentially fuses the neighboring two features progressively with emphasis on the spatial semantics. The network uses the Segformer network with high performance as a backbone in multiple computer vision tasks to enhance the segmentation capability. Experimental results obtained from two publicly available datasets of indoor scenes reveal that the proposed CMPFFNet outperforms existing models in semantic segmentation of indoor scenes of RGB-D images.  
Note to Practitioners-This study introduces a cross-modal and progressive feature fusion network (CMPFFNet) for indoor scene semantic segmentation in RGB-D images. The complementary information of the depth modality is incorporated into the RGB modality in both channel and spatial forms to form a discriminative representation for easy segmentation. A multilevel feature aggregation decoder is proposed to predict the results of semantic segmentation of scenes. The network uses the Segformer network with high performance as a backbone in multiple computer vision tasks to enhance the segmentation capability.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
1545-5955  
eISSN:  
1558-3783  
入藏号:  
WOS:001071992200001  
Indexed Date:  
2023-10-19  
Institution Address:  
Zhejiang Univ Sci & Technol, Sch Informat & Elect Engn, Hangzhou 310023, Peoples R China  
Institution Address:  
Nanyang Technol Univ, Sch Comp Sci & Engn, Singapore 308232, Singapore  
Institution Address:  
Zhejiang Univ, Inst Informat & Commun Engn, Hangzhou 310027, Peoples R China  
Institution:  
Institution:  
Institution:

### 877. Complex gesture pose estimation network fusing multiscale features

摘要: Objective Hand pose estimation aims to identify and localize key points of human hands in images. It has a wide range of applications in computer vision. Hand pose estimation methods can be categorized as depth- or RGB-based methods. Depth-based methods estimate the hand pose by extracting depth features. They require specific devices to constrain the user environment. Scholars use RGB images for hand pose estimation. However, this approach is difficult in an occluded environment. In particular, hand pose estimation based on a single RGB image has low accuracy because of the complexity of the pose, local self-similarity of finger features, and occlusion. Edge information is usually ignored in hand pose estimation. However, this information is important in extracting the information of occluded parts. Moreover, fingertips are small, thereby complicating the recognition of the joints at the fingertips. However, many existing RGB-based gesture estimation methods do not make good use of edge information. A multiscale feature fusion network for monocular vision gesture pose estimation is proposed to address this problem. Method Gesture pictures usually contain complex detailed features. A strong correlation between fingers and joints is present. Therefore, the use of a single feature for hand pose estimation tends to ignore diverse feature information, thereby complicating the accurate extraction of gesture information. Multiscale feature fusion network(MS-FF)aims to estimate the hand pose through a single RGB image. The feature maps of different resolutions are extracted from RGB images through the ResNet50 module. Feature maps are fed into the channel conversion module to learn the dependencies between channels explicitly, thereby enhancing important information and downplaying minor information. The level of feature information depends on the resolution of a feature map. Thus, the global regression module obtains high-resolution feature maps containing semantic information. These maps are separately input in the local optimization module to extract deep information. The Gaussian heatmap of hand joints is obtained to improve the spatial generalization ability of the model. Thus, accurate joint locations can be obtained. We take the feature map with the smallest resolution from the channel conversion module, through which the handedness and relative depth information between the wrist joints are obtained. The above results are combined to estimate the hand pose. Result The PyTorch framework was used for training. The hand image was resized to 256 \* 256 pixels and input to the network. In the experiment, the batch size was set to 16. The network was trained for 20 epochs with an NVIDIA 3090 GPU. The initial learning rate was set to 0.000 1 and reduced by a factor of 10 at the 15th and 17th epochs to optimize the network output. The proposed method achieved better metrics than other methods on different test sets. InterHand2.6M(H+M)was selected as the training set. Compared with the evaluation metrics obtained by InterNet, the mean relative root position error, mean per joint position error of single hand sequences, and mean per joint position error of interacting hand sequences obtained by MS-FF had low errors of 30.92, 11.10, and 15.14, respectively. These values were 5.1%, 8.3%, and 5.8% lower than those obtained by InterNet. We also found that each finger achieved a low error. MS-FF also possesses few model parameters and low computational complexity while improving recognition accuracy. However, the running rate of MS-FF(28 frame/s)is lower than that of InterNet(53 frame/s). The picture shows the hand pose with finger self-occlusion and mutual occlusion of hands. Thus, estimating this interacting hand pose is more difficult than predicting a single hand pose. In the result obtained by our method, the hand joint positions and hand pose estimations are correctly predicted under occlusion.  
摘要:  
目的基于单幅RGB图像的手势姿态估计受手势复杂性、手指特征局部自相似性及遮挡问题的影响,导致手势姿态估计准确率低。为此,提出一种面向单目视觉手势姿态估计的多尺度特征融合网络。方法1)采用ResNet50(50-layer residual network)模块从RGB图像提取不同分辨率特征图,通过通道变换模块显式地学习特征通道间的依赖关系,增强重要的特征通道信息,弱化次要的特征通道信息。2)在全局回归模块中,通过设计节点间的连接方式融合不同分辨率特征图,以便充分利用图像的细节与整体信息。采用局部优化模块继续提取更深层的特征信息,获得手部关节点的高斯热图,以此修正遮挡等原因造成部分关节点回归不准确的问题。3)计算经通道变换模块处理后的最小特征图,通过全局池化和多层感知机处理该特征图以获得手势类别和右手相对于左手的深度。 4)综合以上结果获得最终的手势姿态。结果采用InterHand2.6M和RHD(rendered handpose dataset)数据集训练多尺度特征融合网络,评估指标中根节点的平均误差和关节点的平均误差,均低于同类方法,且在一些复杂和遮挡的场景下鲁棒性更高。在InterHand2.6M数据集上,与InterNet方法相比,本文方法的交互手关节点的平均误差降低5.8%,单手关节点的平均误差降低8.3%,根节点的平均误差降低5.1%。从RHD数据集的测试结果看,与同类方法相比,本文方法在手部关节点的平均误差上获得最小值。结论本文提出的多尺度特征融合网络能够更准确地预测手部关节点位置,适用于复杂手势或遮挡条件下的手势姿态估计(本文方法代码网址:https://github. com/cornersInHeart/hand-pose-esitmation. git)。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-8961  
入藏号:  
CSCD:7560509  
Indexed Date:  
2023-12-23  
Institution Address:  
辽宁工程技术大学电子与信息工程学院;;辽宁工程技术大学电气与控制工程学院, ;;, 葫芦岛;;葫芦岛, ;; 125105;;125105  
Institution Address:  
辽宁工程技术大学电子与信息工程学院, 葫芦岛, 辽宁 125105, 中国  
Institution Address:  
School of Electronic and Information Engineering, Liaoning Technical University;;Faculty of Electrical and Control Engineering, Liaoning Technical University, ;;, Huludao;;Huludao, ;; 125105;;125105  
Institution Address:  
School of Electronic and Information Engineering, Liaoning Technical University, Huludao, Liaoning 125105, China  
Institution:  
辽宁工程技术大学电子与信息工程学院;;辽宁工程技术大学电气与控制工程学院  
Institution:  
辽宁工程技术大学电子与信息工程学院  
Institution:  
School of Electronic and Information Engineering, Liaoning Technical University;;Faculty of Electrical and Control Engineering, Liaoning Technical University  
Institution:  
School of Electronic and Information Engineering, Liaoning Technical University

### 878. MFCF-Gait: Small Silhouette-Sensitive Gait Recognition Algorithm Based

摘要: Gait recognition based on gait silhouette profiles is currently a major approach in the field of gait recognition. In previous studies, models typically used gait silhouette images sized at 64 x 64 pixels as input data. However, in practical applications, cases may arise where silhouette images are smaller than 64 x 64, leading to a loss in detail information and significantly affecting model accuracy. To address these challenges, we propose a gait recognition system named Multi-scale Feature Cross-Fusion Gait (MFCF-Gait). At the input stage of the model, we employ super-resolution algorithms to preprocess the data. During this process, we observed that different super-resolution algorithms applied to larger silhouette images also affect training outcomes. Improved super-resolution algorithms contribute to enhancing model performance. In terms of model architecture, we introduce a multi-scale feature cross-fusion network model. By integrating low-level feature information from higher-resolution images with high-level feature information from lower-resolution images, the model emphasizes smaller-scale details, thereby improving recognition accuracy for smaller silhouette images. The experimental results on the CASIA-B dataset demonstrate significant improvements. On 64 x 64 silhouette images, the accuracies for NM, BG, and CL states reached 96.49%, 91.42%, and 78.24%, respectively. On 32 x 32 silhouette images, the accuracies were 94.23%, 87.68%, and 71.57%, respectively, showing notable enhancements.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001311409600001  
PubMed ID:  
39275411  
Indexed Date:  
2024-09-16  
Institution Address:  
Yunnan Normal Univ, Coll Informat, Kunming 650500, Peoples R China  
Institution Address:  
Engn Res Ctr Comp Vis & Intelligent Control Techno, Dept Educ Yunnan Prov, Kunming 650500, Peoples R China  
Institution:  
Institution:  
Engn Res Ctr Comp Vis & Intelligent Control Techno

### 879. A Knowledge Distillation-Based Ground Feature Classification Network

摘要: As a fundamental task in remote-sensing interpretation, semantic segmentation of remote-sensing images intends to allocate a definite class to each pixel in the image. Fast and efficient semantic segmentation of high-resolution remote-sensing images provides help to capture the real surface covering and plays an essential role in urban planning and dynamic monitoring. However, there are still some limitations in the previous remote-sensing image semantic segmentation model for urban scenes, such as the low weight of small target pixels and the tiny target size leading to the unsatisfactory recognition and segmentation results of the model for small target features. Meanwhile, the deeper and broader feature extraction module in the semantic segmentation network usually leads to more redundant parameters, which takes a lot of computation time. Thus, we propose a lightweight semantic segmentation network based on the knowledge distillation combined with a multiscale pyramidal pooling module and attention mechanism named KD-MSANet, which enhanced the ability to fuse and focus on shallow features. Then, we trained teacher-student models to obtain lightweight network models through a model pruning and distillation framework. Experiments on Vaihingen and Potsdam datasets demonstrated that the network we designed significantly reduces the number of parameters while ensuring almost constant accuracy. Compared with the precompression model, the student model reduced in size by 43.6% and the training efficiency was improved by 22.3%, while the accuracy reached 99.30% of the teacher model.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001138687300004  
Indexed Date:  
2024-02-04  
Institution Address:  
Capital Normal Univ, 3D Informat Collect & Applicat Key Lab Educ Minis, Beijing 100048, Peoples R China  
Institution Address:  
Capital Normal Univ, Coll Resources Environm & Tourism, Beijing State Key Lab Incubat Base Urban Environm, Beijing 100048, Peoples R China  
Institution:  
Institution:

### 880. ABYOLOv4: improved YOLOv4 human object detection based on enhanced

摘要: The purpose of human object detection is to obtain the number of people and their position in images, which is one of the core problems in the field of machine vision. However, the high missing detection rate from small- and medium-sized human bodies due to the large variety of human scale in human object detection tasks still influences the performance of human object detection. To solve the above problem, this paper proposed an improved ASPP\_BiFPN\_YOLOv4 (ABYOLOv4) method to detect human object detection. In detail, Atrous Spatial Pyramid Pooling (ASPP) module was used to replace the original Spatial Pyramid Pooling module to increase the receptive field level of the network and improve the perception ability of multi-scale targets. Then, the original Path Aggregation Network (PANet) multi-scale fusion module was replaced by the self-built bi-layer bidirectional feature pyramid network (Bi-FPN). Meanwhile, a new feature was imported into the proposed model to reuse the mid- and low-level features, which could enhance the ability of the network to express the characteristics of small- and medium-sized targets. Finally, the standard convolution in Bi-FPN was replaced by depth-separable convolution to make the network achieve the balance of accuracy and the number of parameters. To identify the performance of the proposed ABYOLOv4 model, the human object detection experiment is carried out by using the public data set of VOC2007 and VOC2012, the improved YOLOv4 algorithm is 0.5% higher than the original AP algorithm, and the weight file size of the model is reduced by 45.3 M. The experimental results demonstrated that the proposed ABYOLOv4 network has higher accuracy and lower computational cost for human target detection.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
ISSN:  
1687-6180  
入藏号:  
WOS:001136313000001  
Indexed Date:  
2024-01-18  
Institution Address:  
Xian Univ Technol, Sch Mech & Precis Instrument Engn, Xian 710048, Peoples R China  
Institution Address:  
Xian Peoples Hosp, Xian, Peoples R China  
Institution:  
Institution:  
Xian Peoples Hosp

### 881. FMR-YOLO: Infrared Ship Rotating Target Detection Based on Synthetic Fog

摘要: Infrared ship detection has important application value for ensuring navigation safety and real-time monitoring of the sea surface. It is also of great significance in marine intelligent defense and has become an important research branch in the field of computer vision. Affected by the weather at sea and the limitations of infrared cameras, infrared ship images often have the problems of small targets being submerged by noise and low information entropy, which bring great challenges to infrared ship detection. In this article, an infrared ship rotating target detection algorithm FMR-YOLO based on synthetic fog and multiscale weighted fusion is proposed. Our algorithm first corrects the noisy labels of the original dataset due to misclassification and constructs an infrared ship dataset (ISD) containing different concentrations of haze through an improved dark channel prior (DCP) algorithm. Second, in order to avoid the loss of small target features and information as the network deepens, a weighted feature pyramid network (FPN) based on dilated convolution (DWFPN) is proposed. DWFPN weights the fusion of features at different levels based on the attention mechanism to achieve high-quality information interaction. Finally, in view of the large aspect ratio and arbitrary direction of the ship target, rotation detection is introduced to obtain more accurate detection boxes and ship navigation direction information. The experimental results show that compared with the standard YOLOv7, the improved algorithm achieves a mean average accuracy (mAP) of 92.7%, and the recall rate and precision rate are improved by 2.3% and 3%, respectively. Our code and R-ISD dataset are available at: https://github.com/denghuimin1/FMR-YOLO.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
22  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
22  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001132683400173  
Indexed Date:  
2024-02-02  
Institution Address:  
Shanghai Maritime Univ, Coll Informat Engn, Shanghai 201306, Peoples R China  
Institution:

### 882. A Wildfire Smoke Detection System Using Unmanned Aerial Vehicle Images

摘要: Wildfire is one of the most significant dangers and the most serious natural catastrophe, endangering forest resources, animal life, and the human economy. Recent years have witnessed a rise in wildfire incidents. The two main factors are persistent human interference with the natural environment and global warming. Early detection of fire ignition from initial smoke can help firefighters react to such blazes before they become difficult to handle. Previous deep-learning approaches for wildfire smoke detection have been hampered by small or untrustworthy datasets, making it challenging to extrapolate the performances to real-world scenarios. In this study, we propose an early wildfire smoke detection system using unmanned aerial vehicle (UAV) images based on an improved YOLOv5. First, we curated a 6000-wildfire image dataset using existing UAV images. Second, we optimized the anchor box clustering using the K-mean++ technique to reduce classification errors. Then, we improved the network's backbone using a spatial pyramid pooling fast-plus layer to concentrate small-sized wildfire smoke regions. Third, a bidirectional feature pyramid network was applied to obtain a more accessible and faster multi-scale feature fusion. Finally, network pruning and transfer learning approaches were implemented to refine the network architecture and detection speed, and correctly identify small-scale wildfire smoke areas. The experimental results proved that the proposed method achieved an average precision of 73.6% and outperformed other one- and two-stage object detectors on a custom image dataset.  
在 BIOSIS Citation Index 中的被引频次:  
8  
Web of Science 核心合集中的 "被引频次":  
52  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
52  
eISSN:  
1424-8220  
入藏号:  
WOS:000896475700001  
PubMed ID:  
36502081  
Indexed Date:  
2022-12-24  
Institution Address:  
Gachon Univ, Dept Comp Engn, Seongnam Si 13120, South Korea  
Institution:

### 883. Segmentation network for metastatic lymph nodes of head and neck tumors

摘要: Head and neck tumors are prevalent malignant tumors in China, with prognosis significantly influenced by cervical lymph node metastasis. In medical practice, magnetic resonance imaging (MRI) is employed to identify metastatic lymph nodes. However, MRI images often suffer from blurred edges and low contrast between the lesion and surrounding tissue. This paper introduces a segmentation network tailored for metastatic lymph nodes in head and neck tumors. Initially, a cross-layer and cross-field attention module is developed, integrating features from both deep and shallow layers to enhance the shape representation of metastatic lymph nodes through a self-attention mechanism. This module improves contextual semantic understanding across different receptive fields, allowing for pixel-level fusion of shallow and deep feature maps, thereby enhancing the morphological details of metastatic lymphatic nodes. Subsequently, a multi-scale feature fusion module is designed to amalgamate features across various scales in the feature pyramid, enriching the morphological details of the lymph nodes. Furthermore, an enhanced attention prediction head module is implemented, combining parallel self-attention and gate channel transformation to accentuate the lesion area and refine its boundaries on the feature map. The network's effectiveness is confirmed using a clinical dataset of lymph node metastasis medical images. The performance metrics, APdet, APseg, ARdet, ARseg, mAPdet, and mAPseg for lymph node metastasis lesion segmentation are 74.88%, 74.12%, 63.11%, 62.28%, 74.64%, and 74.04%, respectively. This network provides precise detection and segmentation of lymph node metastasis lesions, offering significant benefits for lymph node diagnosis.  
摘要:  
头颈部肿瘤是我国常见的恶性肿瘤,其预后主要受颈部淋巴结转移的影响,医学上通过核磁共振成像技术对转移性淋巴结成像后再进行诊断,然而,核磁共振技术成像存在病灶形态信息丢失,病灶区域对比度低和病灶边界模糊的问题。针对这些问题,提出面向头颈部肿瘤转移性淋巴结分割网络协助医生进行诊断。首先,设计跨层跨视野注意力模块,其接收深浅层的特征信息后利用自注意力机制分别突显深浅层的转移淋巴形状,通过不同感受野的深层特征图学习到更好的语义上下文特征,将浅层特征图与深层特征图逐像素融合,增强转移淋巴病灶区域的形态信息。其次,设计多尺度特征融合模块,在特征金字塔的初始位置融合不同尺度的特征图,丰富转移淋巴病灶区域的形态信息。然后,设计增强注意力预测头模块,通过对预测前的特征图使用并行的自注意力与门控通道转换模块,凸出病灶区域,细化病灶边界。最后,使用临床淋巴结转移医学图像数据集验证网络的有效性。实验结果表明,所提网络对于淋巴结转移病灶分割的APdet,APseg,ARdet,ARseg,mAPdet和mAPseg分别为74.88%,74.12%,63.11%,62.28%,74.64%和74.04%。该网络实现对淋巴结转移病灶区域的精确检测分割,对辅助淋巴结诊断具有积极意义。  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1004-924X  
入藏号:  
CSCD:7720472  
Indexed Date:  
2024-07-26  
Institution Address:  
北方民族大学计算机科学与工程学院;;北方民族大学, ;;图像图形智能处理国家民委重点实验室, 银川;;银川, 宁夏;;宁夏 750021;;750021, 中国  
Institution Address:  
宁夏医科大学口腔学院, 银川, 宁夏 750004, 中国  
Institution Address:  
College of Computer Science and Engineering, North Minzu University;;North Minzu University, ;;Key Laboratory of Image and Graphics Intelligent Processing of State Ethnic Affairs Commission, Yinchuan;;Yinchuan, ;; 750021;;750021  
Institution Address:  
College of Oral Cavity, Ningxia Medical University, Yinchuan, Ningxia 750004, China  
Institution:  
北方民族大学计算机科学与工程学院;;北方民族大学  
Institution:  
宁夏医科大学口腔学院  
Institution:  
College of Computer Science and Engineering, North Minzu University;;North Minzu University  
Institution:  
College of Oral Cavity, Ningxia Medical University

### 884. A lightweight white blood cells detection network based on CenterNet and

摘要: White blood cells (WBCs) detection is significant to the diagnosis of many diseases. However, the detection accuracy can be influenced by the significant differences in color, size and morphology of WBCs in the images. In this paper, an improved CenterNet with smaller model size, fewer parameters, and lower computational complexity is proposed for the WBCs accurate detection. Firstly, the lightweight backbone GhostNetv2 is selected to reduce the model size, parameters and computational complexity of the network. Then, a feature pyramid network with the efficient channel attention (ECA) attention mechanism and the multi-scale feature extraction modules is constructed to enhance the capability of feature extraction and focus more on WBCs. Finally, the heatmap loss function is modified by proposing an improved mean squared error loss function to enhance the fitting ability between the predicted values and the ground truth of the heatmap. Experimental results show that the model size of the proposed lightweight CenterNet is only 19.9 MB, and the mAP.5 is 97.36%. The model size is reduced by 84% while the mAP.5 and FPS are increased by 0.7% and 10.4 compared to the original CenterNet. Moreover, the detection accuracy of the proposed lightweight CenterNet is comparable to the existing mainstream networks and its detection performance on different datasets is good, while the model size, parameters and computational complexity of the network is significantly reduced, and can be used for WBCs detection effectively.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001207475100001  
Indexed Date:  
2024-04-28  
Institution Address:  
Hunan Univ Sci & Technol, Sch Informat & Elect Engn, Xiangtan, Peoples R China  
Institution:

### 885. PIS-Net: Efficient Medical Image Segmentation Network with Multivariate

摘要: Recently, with more portable diagnostic devices being moved to people anywhere, point-of-care (PoC) imaging has become more convenient and more popular than the traditional "bed imaging". Instant image segmentation, as an important technology of computer vision, is receiving more and more attention in PoC diagnosis. However, the image distortion caused by image preprocessing and the low resolution of medical images extracted by PoC devices are urgent problems that need to be solved. Moreover, more efficient feature representation is necessary in the design of instant image segmentation. In this paper, a new feature representation considering the relationships among local features with minimal parameters and a lower computational complexity is proposed. Since a feature window sliding along a diagonal can capture more pluralistic features, a Diagonal-Axial Multi-Layer Perceptron is designed to obtain the global correlation among local features for a more comprehensive feature representation. Additionally, a new multi-scale feature fusion is proposed to integrate nonlinear features with linear ones to obtain a more precise feature representation. Richer features are figured out. In order to improve the generalization of the models, a dynamic residual spatial pyramid pooling based on various receptive fields is constructed according to different sizes of images, which alleviates the influence of image distortion. The experimental results show that the proposed strategy has better performance on instant image segmentation. Notably, it yields an average improvement of 1.31% in Dice than existing strategies on the BUSI, ISIC2018 and MoNuSeg datasets.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1099-4300  
入藏号:  
WOS:001209996600001  
PubMed ID:  
38667838  
Indexed Date:  
2024-05-05  
Institution Address:  
Xinjiang Univ, Sch Comp Sci & Technol, Urumqi 830046, Peoples R China  
Institution:

### 886. A Fine-Grained Detection Network Model for Soldier Targets Adopting

摘要: Owing to its ability to provide more accurate and detailed battlefield situational information, fine-grained detection research on soldier targets is of significant importance for military decision-making and firepower threat assessment. To address the issues of low detection accuracy and inaccurate classification in the fine-grained detection of soldier targets, we propose a fine-gain soldier target detection model based on the improved YOLOv8 (You Only Look Once v8). First, we developed a multi-branch feature fusion module to effectively fuse multi-scale feature information and used a dynamic deformable attention mechanism to help the detection model focus on key areas in deep-level features. Second, we proposed a decoupled lightweight dynamic head to extract the position and category information of soldier targets separately, effectively solving the problem of misclassification of soldier targets' attack actions under different poses. Finally, we used the Inner Minimum Points Distance Intersection over Union (Inner-MPDIoU) to further improve the convergence speed and accuracy of the network model. The proposed improvements are evaluated through comparative experiments conducted in published twenty-six test groups, and the effectiveness of the proposed method is demonstrated. Compared with the original model, our method achieved a detection precision of 78.9%, a 6.91% improvement; the mAP@50 (mean Average Precision at 50) was 79.6%, a 3.51% increase; and an mAP@50-95 of 63.8%, a gain of 5.28%. The proposed method achieves high precision and recall while reducing the computational complexity of the model, thereby enhancing its efficiency and robustness for fine-grained soldier target detection.  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001293998800001  
Indexed Date:  
2024-08-23  
Institution Address:  
Beijing Inst Technol, Sch Mech Engn, Beijing 100081, Peoples R China  
Institution:

### 887. UN-YOLOv5s: A UAV-Based Aerial Photography Detection Algorithm

摘要: With the progress of science and technology, artificial intelligence is widely used in various disciplines and has produced amazing results. The research of the target detection algorithm has significantly improved the performance and role of unmanned aerial vehicles (UAVs), and plays an irreplaceable role in preventing forest fires, evacuating crowded people, surveying and rescuing explorers. At this stage, the target detection algorithm deployed in UAVs has been applied to production and life, but making the detection accuracy higher and better adaptability is still the motivation for researchers to continue to study. In aerial images, due to the high shooting height, small size, low resolution and few features, it is difficult to be detected by conventional target detection algorithms. In this paper, the UN-YOLOv5s algorithm can solve the difficult problem of small target detection excellently. The more accurate small target detection (MASD) mechanism is used to greatly improve the detection accuracy of small and medium targets, The multi-scale feature fusion (MCF) path is combined to fuse the semantic information and location information of the image to improve the expression ability of the novel model. The new convolution SimAM residual (CSR) module is introduced to make the network more stable and focused. On the VisDrone dataset, the mean average precision (mAP) of UAV necessity you only look once v5s(UN-YOLOv5s) is 8.4% higher than that of the original algorithm. Compared with the same version, YOLOv5l, the mAP is increased by 2.2%, and the Giga Floating-point Operations Per Second (GFLOPs) is reduced by 65.3%. Compared with the same series of YOLOv3, the mAP is increased by 1.8%, and GFLOPs is reduced by 75.8%. Compared with the same series of YOLOv8s, the detection accuracy of the mAP is improved by 1.1%.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
被引频次合计:  
10  
eISSN:  
1424-8220  
入藏号:  
WOS:001028469200001  
PubMed ID:  
37447757  
Indexed Date:  
2023-07-28  
Institution Address:  
Qilu Univ Technol, Shandong Acad Sci, Sch Informat & Automat Engn, Jinan 250353, Peoples R China  
Institution:

### 888. AMFEF-DETR: An End-to-End Adaptive Multi-Scale Feature Extraction and

摘要: To address the challenge of low detection accuracy and slow detection speed in unmanned aerial vehicle (UAV) aerial images target detection tasks, caused by factors such as complex ground environments, varying UAV flight altitudes and angles, and changes in lighting conditions, this study proposes an end-to-end adaptive multi-scale feature extraction and fusion detection network, named AMFEF-DETR. Specifically, to extract target features from complex backgrounds more accurately, we propose an adaptive backbone network, FADC-ResNet, which dynamically adjusts dilation rates and performs adaptive frequency awareness. This enables the convolutional kernels to effectively adapt to varying scales of ground targets, capturing more details while expanding the receptive field. We also propose a HiLo attention-based intra-scale feature interaction (HLIFI) module to handle high-level features from the backbone. This module uses dual-pathway encoding of high and low frequencies to enhance the focus on the details of dense small targets while reducing noise interference. Additionally, the bidirectional adaptive feature pyramid network (BAFPN) is proposed for cross-scale feature fusion, integrating semantic information and enhancing adaptability. The Inner-Shape-IoU loss function, designed to focus on bounding box shapes and incorporate auxiliary boxes, is introduced to accelerate convergence and improve regression accuracy. When evaluated on the VisDrone dataset, the AMFEF-DETR demonstrated improvements of 4.02% and 16.71% in mAP50 and FPS, respectively, compared to the RT-DETR. Additionally, the AMFEF-DETR model exhibited strong robustness, achieving mAP50 values 2.68% and 3.75% higher than the RT-DETR and YOLOv10, respectively, on the HIT-UAV dataset.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
4  
eISSN:  
2504-446X  
入藏号:  
WOS:001342581400001  
Indexed Date:  
2024-11-06  
Institution Address:  
Minzu Univ China, Key Lab Ethn Language Intelligent Anal & Secur Gov, Beijing 100081, Peoples R China  
Institution Address:  
Minzu Univ China, Sch Informat Engn, Beijing 100081, Peoples R China  
Institution:  
Minzu Univ China  
Institution:

### 889. ALDNet: a lightweight and efficient drone detection network

摘要: With the growing use of unmanned aerial vehicles (UAVs), it poses a potential threat to public safety, so drone detection has become an increasingly important research direction. However, the irregular shooting angles, deformations, significant target scale variations, and complex backgrounds in UAV imagery pose challenges for existing object detection models. To address these issues, we propose ALDNet, a precise and lightweight UAV image object detection network built on an improved RT-DETR model. First, we introduce the position accurate continuous convolution (PAC), which dynamically adjusts convolution kernel coordinates, allowing flexibility in detecting objects of varied shapes. PAC improve the ability of backbone network to capture objects with different shapes. Second, we present an enhanced multi-scale feature fusion approach that integrates spatial information more effectively, reducing the loss of fine-grained details during fusion. Additionally, we propose the Inner-Generalized Intersection over Union loss function, which reduces interference and accelerates the convergence of bounding box optimisation, thus improving the overall detection performance. We conducted experiments on three UAV image datasets, Det-Fly, ARD-MAV and VisDrone, and compared the ALDNet results with the RT-DETR model. The experimental results show that ALDNet achieved 98.3%, 97.4% and 48.1% mAP50 values on Det-Fly, ARD-MAV and VisDrone, which are 1.5%, 0.4% and 1.9% higher than RT-DETR, respectively, and the number of parameters and computation amount of ALDNet are 25% and 14% less than the baseline, respectively. Experimental results show that ALDNet achieves greater efficiency and effectiveness with less computational power.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001400562200001  
Indexed Date:  
2025-01-25  
Institution Address:  
Shantou Univ, Coll Math & Comp Sci, Shantou 515063, Peoples R China  
Institution:

### 890. LMFUNet: A Lightweight Multi-fusion UNet Based on Spiking Neural Systems

摘要: Skin lesion segmentation is critical in medical image processing, but the segmentation task faces numerous challenges due to the differences in size, color, shape, and texture of skin lesions between patients, as well as the blurring of the boundary between lesions and normal skin. While many models improve segmentation performance by introducing complex modules, their computational resource constraints make applications limited in a clinical setting. To cope with this problem, we propose a lightweight multi-fusion network (LMFUNet) with parameters of only 0.100M and GFLOPs of 0.106. LMFUNet uses an Efficient Multi-scale Feature Extraction block (EMFE) in deep stages, which uses grouping of features by convolution with different dilation rates to reduce model complexity and effectively capture multi-scale features. By using the Multi-level Feature Fusion module (MFF) in skip connections, different levels of information are combined step by step, realizing the first step of fusion of low-level details and high-level contextual information, which helps to accurately localize the lesion area. After MFF, we designed the Spatial-channel Fusion module (SCF). This module further optimizes the fusion of feature information output from MFF in terms of spatial and channel dimensions, significantly enhancing the recognition of lesion boundaries. We did in-depth tests on three public datasets that are typical of others: ISIC2017, ISIC2018 and PH2 datasets. LMFUNet demonstrates superior performance because it excels at segmentation and doesn't require a lot of computing power.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001386426100001  
Indexed Date:  
2025-01-05  
Institution Address:  
Xihua Univ, Sch Comp & Software Engn, Chengdu 610039, Peoples R China  
Institution Address:  
Xihua Univ, Sch Elect Engn & Elect Informat, Chengdu 610039, Peoples R China  
Institution:  
Institution:

### 891. Subtask Attention Based Object Detection in Remote Sensing Images

摘要: Object detection in remote sensing images (RSIs) is one of the basic tasks in the field of remote sensing image automatic interpretation. In recent years, the deep object detection frameworks of natural scene images (NSIs) have been introduced into object detection on RSIs, and the detection performance has improved significantly because of the powerful feature representation. However, there are still many challenges concerning the particularities of remote sensing objects. One of the main challenges is the missed detection of small objects which have less than five percent of the pixels of the big objects. Generally, the existing algorithms choose to deal with this problem by multi-scale feature fusion based on a feature pyramid. However, the benefits of this strategy are limited, considering that the location of small objects in the feature map will disappear when the detection task is processed at the end of the network. In this study, we propose a subtask attention network (StAN), which handles the detection task directly on the shallow layer of the network. First, StAN contains one shared feature branch and two subtask attention branches of a semantic auxiliary subtask and a detection subtask based on the multi-task attention network (MTAN). Second, the detection branch uses only low-level features considering small objects. Third, the attention map guidance mechanism is put forward to optimize the network for keeping the identification ability. Fourth, the multi-dimensional sampling module (MdS), global multi-view channel weights (GMulW) and target-guided pixel attention (TPA) are designed for further improvement of the detection accuracy in complex scenes. The experimental results on the NWPU VHR-10 dataset and DOTA dataset demonstrated that the proposed algorithm achieved the SOTA performance, and the missed detection of small objects decreased. On the other hand, ablation experiments also proved the effects of MdS, GMulW and TPA.  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
eISSN:  
2072-4292  
入藏号:  
WOS:000662614900001  
Indexed Date:  
2021-07-07  
Institution Address:  
Huazhong Univ Sci & Technol, Sch Artificial Intelligence & Automat, Natl Key Lab Sci & Technol Multispectral Informat, Wuhan 430074, Peoples R China  
Institution Address:  
Shenzhen Huazhong Univ Sci & Technol, Res Inst, Shenzhen 518000, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Remote Sensing & Informat Engn, Wuhan 430079, Peoples R China  
Institution:  
Institution:  
Institution:

### 892. Three-Dimensional Millimeter-Wave Object Detector Based on the

摘要: Millimeter-wave (MMW) point clouds, characterized by their low resolution and high noise, limit the detection accuracy of point-based IA-SSD method due to the inadequate consideration of contextual information in MMW scenarios. Therefore, this paper proposes a three-dimensional (3D) MMW object detector, greatly augmenting the detection performance of the baseline model IA-SSD by the integration of the local-global context information. Central to our approach is the implementation of a multi-scale feature aggregation (MFA) module in the encoder stage of IA-SSD, which utilizes a self-attention mechanism to apprehend local contextual distinctions. This module is further applied to the centroid aggregation stage to enhance the capture of local context from foreground points. Complementarily, a global feature fusion module is devised to combine global contextual insights, drawing upon the localized information delineated by the MFA modules. This integrated framework significantly diminishes the false detection rate while concurrently elevating the detection precision for occluded objects. Relative to the IA-SSD baseline, the empirical evaluations validate the efficiency of our proposed model, demonstrating marked decreases in false positives and false negatives. Specifically, there is a 2.78% and 7.39% improvement in AP\_R40\_0.25 and AP\_R40\_0.5, respectively. When the intersection-over-union threshold is set as 0.25 and 0.5, the corresponding recall rate increases by 2.13% and 6.2%, respectively. Moreover, the inference speed reaches 32.3 frames per second(FPS), only a slight decrease of 2.9 FPS compared to the baseline model. These results demonstrate that the proposed detector significantly enhances detection performance without compromising on speed, marking a considerable advancement in the domain of 3D MMW object detection.  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001320512300001  
Indexed Date:  
2024-10-01  
Institution Address:  
Univ Shanghai Sci & Technol, Inst Biomed Engn, Shanghai 200093, Peoples R China  
Institution Address:  
Univ Shanghai Sci & Technol, Terahertz Technol Innovat Res Inst, Shanghai 200093, Peoples R China  
Institution:  
Institution:

### 893. Real-Time Ground-Level Building Damage Detection Based on Lightweight

摘要: Real-time building damage detection effectively improves the timeliness of post-earthquake assessments. In recent years, terrestrial images from smartphones or cameras have become a rich source of disaster information that may be useful in assessing building damage at a lower cost. In this study, we present an efficient method of building damage detection based on terrestrial images in combination with an improved YOLOv5. We compiled a Ground-level Detection in Building Damage Assessment (GDBDA) dataset consisting of terrestrial images with annotations of damage types, including debris, collapse, spalling, and cracks. A lightweight and accurate YOLOv5 (LA-YOLOv5) model was used to optimize the detection efficiency and accuracy. In particular, a lightweight Ghost bottleneck was added to the backbone and neck modules of the YOLOv5 model, with the aim to reduce the model size. A Convolutional Block Attention Module (CBAM) was added to the backbone module to enhance the damage recognition effect. In addition, regarding the scale difference of building damage, the Bi-Directional Feature Pyramid Network (Bi-FPN) for multi-scale feature fusion was used in the neck module to aggregate features with different damage types. Moreover, depthwise separable convolution (DSCONV) was used in the neck module to further compress the parameters. Based on our GDBDA dataset, the proposed method not only achieved detection accuracy above 90% for different damage targets, but also had the smallest weight size and fastest detection speed, which improved by about 64% and 24%, respectively. The model performed well on datasets from different regions. The overall results indicate that the proposed model realizes rapid and accurate damage detection, and meets the requirement of lightweight embedding in the future.  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
22  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
23  
eISSN:  
2072-4292  
入藏号:  
WOS:000818305600001  
Indexed Date:  
2022-07-06  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & R, Wuhan 430079, Peoples R China  
Institution Address:  
Tianjin Inst Surveying & Mapping Co Ltd, 9 Changling Rd, Tianjin 300381, Peoples R China  
Institution:  
Institution:  
Tianjin Inst Surveying & Mapping Co Ltd

### 894. Embedded Sensing System for Recognizing Citrus Flowers Using Cascaded

摘要: Florescence information monitoring is essential for strengthening orchard management activities, such as flower thinning, fruit protection, and pest control. A lightweight object recognition model using cascade fusion YOLOv4-CF is proposed, which recognizes multi-type objects in their natural environments, such as citrus buds, citrus flowers, and gray mold. The proposed model has an excellent representation capability with an improved cascade fusion network and a multi-scale feature fusion block. Moreover, separable deep convolution blocks were employed to enhance object feature information and reduce model computation. Further, channel shuffling was used to address missing recognition in the dense distribution of object groups. Finally, an embedded sensing system for recognizing citrus flowers was designed by quantitatively applying the proposed YOLOv4-CF model to an FPGA platform. The mAP@.5 of citrus buds, citrus flowers, and gray mold obtained on the server using the proposed YOLOv4-CF model was 95.03%, and the model size of YOLOv4-CF + FPGA was 5.96 MB, which was 74.57% less than the YOLOv4-CF model. The FPGA side had a frame rate of 30 FPS; thus, the embedded sensing system could meet the demands of florescence information in real-time monitoring.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
7  
eISSN:  
1424-8220  
入藏号:  
WOS:000760073600001  
PubMed ID:  
35161998  
Indexed Date:  
2022-03-06  
Institution Address:  
South China Agr Univ, Coll Elect Engn, Coll Artificial Intelligence, Guangzhou 510642, Peoples R China  
Institution Address:  
Guangdong Lab Lingnan Modern Agr, Guangzhou 510642, Peoples R China  
Institution Address:  
Pazhou Lab, Guangzhou 510330, Peoples R China  
Institution Address:  
China Agr Res Syst MOF & MARA, Div Citrus Machinery, Guangzhou 510642, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
China Agr Res Syst MOF & MARA

### 895. EMR-HRNet: A Multi-Scale Feature Fusion Network for Landslide

摘要: Landslides constitute a significant hazard to human life, safety and natural resources. Traditional landslide investigation methods demand considerable human effort and expertise. To address this issue, this study introduces an innovative landslide segmentation framework, EMR-HRNet, aimed at enhancing accuracy. Initially, a novel data augmentation technique, CenterRep, is proposed, not only augmenting the training dataset but also enabling the model to more effectively capture the intricate features of landslides. Furthermore, this paper integrates a RefConv and Multi-Dconv Head Transposed Attention (RMA) feature pyramid structure into the HRNet model, augmenting the model's capacity for semantic recognition and expression at various levels. Last, the incorporation of the Dilated Efficient Multi-Scale Attention (DEMA) block substantially widens the model's receptive field, bolstering its capability to discern local features. Rigorous evaluations on the Bijie dataset and the Sichuan and surrounding area dataset demonstrate that EMR-HRNet outperforms other advanced semantic segmentation models, achieving mIoU scores of 81.70% and 71.68%, respectively. Additionally, ablation studies conducted across the comprehensive dataset further corroborate the enhancements' efficacy. The results indicate that EMR-HRNet excels in processing satellite and UAV remote sensing imagery, showcasing its significant potential in multi-source optical remote sensing for landslide segmentation.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
eISSN:  
1424-8220  
入藏号:  
WOS:001245687900001  
PubMed ID:  
38894469  
Indexed Date:  
2024-06-19  
Institution Address:  
Jiangxi Univ Sci & Technol, Sch Civil Engn & Surveying & Mapping Engn, Ganzhou 341000, Peoples R China  
Institution:

### 896. Highly Accurate and Lightweight Detection Model of Apple Leaf Diseases

摘要: To mitigate problems concerning small-sized spots on apple leaves and the difficulties associated with the accurate detection of spot targets exacerbated by the complex backgrounds of orchards, this research used alternaria leaf spots, rust, brown spots, gray spots, and frog eye leaf spots on apple leaves as the research object and proposed the use of a high-accuracy detection model YOLOv5-Res (YOLOv5-Resblock) and lightweight detection model YOLOv5-Res4 (YOLOv5-Resblock-C4). Firstly, a multiscale feature extraction module, ResBlock (residual block), was designed by combining the Inception multi-branch structure and ResNet residual idea. Secondly, a lightweight feature fusion module C4 (CSP Bottleneck with four convolutions) was designed to reduce the number of model parameters while improving the detection ability of small targets. Finally, a parameter-streamlining strategy based on an optimized model architecture was proposed. The experimental results show that the performance of the YOLOv5-Res model and YOLOv5-Res4 model is significantly improved, with the mAP0.5 values increasing by 2.8% and 2.2% compared to the YOLOv5s model and YOLOv5n model, respectively. The sizes of the YOLOv5-Res model and YOLOv5-Res4 model are only 10.8 MB and 2.4 MB, and the model parameter counts are reduced by 22% and 38.3% compared to the YOLOv5s model and YOLOv5n model.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2073-4395  
入藏号:  
WOS:001254434400001  
Indexed Date:  
2024-07-02  
Institution Address:  
Yanshan Univ, Sch Mech Engn, Qinhuangdao 066004, Peoples R China  
Institution Address:  
Tangshan Jidong Equipment Engn, Tangshan 063200, Peoples R China  
Institution Address:  
Hebei Univ Engn, Sch Mech & Equipment Engn, Handan 056038, Peoples R China  
Institution:  
Institution:  
Tangshan Jidong Equipment Engn  
Institution:

### 897. Automatic crack defect detection via multiscale feature aggregation and

摘要: In this paper, a multi-scale feature aggregation and adaptive fusion network, is proposed for automatic and accurate pavement crack defect segmentation. Specifically, faced with the linear characteristic of pavement crack defects, a multiple-dimension attention (MDA) module is proposed to effectively capture long-range correlation from three directions, including space, width and height, and help identify the pavement crack defect boundaries. On this basis, a multi-scale skip connection (MSK) module is proposed, which can effectively utilize the feature information from multiple receptive fields to support accurate feature reconstruction in the decoding stage. Furthermore, a multi-scale attention fusion (MSAF) module is proposed to realize effective multi-scale feature representation and aggregation. Finally, an adaptive weight fusion (AWL) module proposed to dynamically fuse the output features across different network layers for accurate multi-scale crack defect segmentation. Experiments indicate that proposed network is superior to other mainstream segmentation networks on pixelwise crack defect detection task.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0926-5805  
eISSN:  
1872-7891  
入藏号:  
WOS:001393931200001  
Indexed Date:  
2025-01-17  
Institution Address:  
Zhengzhou Univ, Sch Elect & Informat Engn, Zhengzhou 450001, Peoples R China  
Institution:

### 898. MGU-Net: a multiscale gate attention encoder-decoder network for medical

摘要: Medical image segmentation, the prerequisite of numerous clinical needs, has been significantly prospered from recent advances in encoder-decoder networks. However, uneven reflection of human organs and the subject's tremor and movement cause blurred edges in the image, which is difficult to segment. Hence and more details and context information are needed to resolve this problem. Most of the existing Unet-like architectures do not take into account the multiscale characteristics of medical images and do not make full use of the spatial information and channel information of feature maps, resulting in the loss of detail information. This paper proposes a Multiscale Gate Attention (MGU-Net) encoder-decoder network. Firstly, we use multiscale blocks to focus on the fusion of contextual information. Besides, we use two gate attention to deploy more detailed information. On three different public datasets, compared with other State-of-the-Art (SOTA) methods, the proposed method achieves an improvement.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0952-8091  
eISSN:  
1741-5047  
入藏号:  
WOS:001035447200001  
Indexed Date:  
2023-08-15  
Institution Address:  
Shaanxi Univ Sci & Technol, Shaanxi Joint Lab Artificial Intelligence, Xian, Shaanxi, Peoples R China  
Institution Address:  
Shaanxi Univ Sci & Technol, Coll Elect Informat & Artificial Intelligence, Xian, Shaanxi, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Affiliated Hosp 1, Dept Geriatr Surg, Xian, Shaanxi, Peoples R China  
Institution:  
Institution:  
Institution:

### 899. Improved Ship Object Detection in Low-Illumination Environments Using

摘要: Video-based ship object detection has long been a popular research issue that has received attention in the water transportation industry. However, in low-illumination environments, such as at night or in fog, the water environment has a complex variety of light sources, video surveillance images are often accompanied by noise, and information on the details of objects in images is worsened. These problems cause high rates of false detection and missed detection when performing object detection for ships in low-illumination environments. Thus, this paper takes the detection of ship objects in low-illumination environments at night as the research object. The technical difficulties faced by object detection algorithms in low-illumination environments are analyzed, and a dataset of ship images is constructed by collecting images of ships (in the Nanjing section of Yangtze River in China) in low-illumination environments. In view of the outstanding performance of the RetinaNet model in general object detection, a new multiscale feature fusion network structure for a feature extraction module is proposed based on the same network architecture, in such a way that the extraction of more potential feature information from low-illumination images can be realized. In line with the feature detection network, the regression and classification detection network for anchor boxes is improved by means of the attention mechanism, guiding the network structure in the detection of object features. Moreover, the design and optimization of the augmentation of multiple random images and prior bounding boxes in the training process are also carried out. Finally, on the basis of experimental validation analysis, the optimized detection model was able to improve ship detection accuracy by 3.7% with a limited decrease in FPS (frames per second), and has better results in application.  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
eISSN:  
2077-1312  
入藏号:  
WOS:000902436500001  
Indexed Date:  
2023-01-15  
Institution Address:  
Dalian Maritime Univ, Key Lab Marine Simulat & Control, Dalian 116026, Peoples R China  
Institution Address:  
MOT, Tianjin Res Inst Water Transport Engn, Tianjin 300456, Peoples R China  
Institution:  
Institution:

### 900. HE-YOLO: Aerial Target Detection Based On Improved YOLOv3

摘要: Aerial image-based target detection has problems such as low accuracy in multiscale target detection situations, slow detection speed, missed targets and falsely detected targets. To solve this problem, this paper proposes a detection algorithm based on the improved You Only Look Once (YOLO)v3 network architecture from the perspective of model efficiency and applies it to multiscale image-based target detection. First, the K-means clustering algorithm is used to cluster an aerial dataset and optimize the anchor frame parameters of the network to improve the effectiveness of target detection. Second, the feature extraction method of the algorithm is improved, and a feature fusion method is used to establish a multiscale (large-, medium-, and small-scale) prediction layer, which mitigates the problem of small target information loss in deep networks and improves the detection accuracy of the algorithm. Finally, label regularization processing is performed on the predicted value, the generalized intersection over union (GIoU) is used as the bounding box regression loss function, and the focal loss function is integrated into the bounding box confidence loss function, which not only improves the target detection accuracy but also effectively reduces the false detection rate and missed target rate of the algorithm. An experimental comparison on the RSOD and NWPU VHR-10 aerial datasets shows that the detection effect of high-efficiency YOLO (HE-YOLO) is significantly improved compared with that of YOLOv3, and the average detection accuracies are increased by 8.92% and 7.79% on the two datasets, respectively. The algorithm not only shows better detection performance for multiscale targets but also reduces the missed target rate and false detection rate and has good robustness and generalizability.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0218-0014  
eISSN:  
1793-6381  
入藏号:  
WOS:000731966600001  
Indexed Date:  
2021-12-30  
Institution Address:  
Xinjiang Univ, Sch Software, Key Lab Signal Detect & Proc Xinjiang Uygur Auton, Urumqi 830091, Peoples R China  
Institution:

### 901. Frequency-aware Camouflaged Object Detection

摘要: Camouflaged object detection (COD) is important as it has various potential applications. Unlike salient object detection (SOD), which tries to identify visually salient objects, COD tries to detect objects that are visually very similar to the surrounding background. We observe that recent COD methods try to fuse features from different levels using some context aggregation strategies originally developed for SOD. Such an approach, however, may not be appropriate for COD as these existing context aggregation strategies are good at detecting distinctive objects while weakening the features from less discriminative objects. To address this problem, we propose in this article to exploit frequency learning to suppress the confusing high-frequency texture information, to help separate camouflaged objects from their surrounding background, and a frequency-based method, called FBNet, for camouflaged object detection. Specifically, we design a frequency-aware context aggregation (FACA) module to suppress high-frequency information and aggregate multi-scale features from a frequency perspective, an adaptive frequency attention (AFA) module to enhance the features of the learned important frequency components, and a gradient-weighted loss function to guide the proposed method to pay more attention to contour details. Experimental results show that our model outperforms relevant state-of-the-art methods.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
18  
被引频次合计:  
19  
ISSN:  
1551-6857  
eISSN:  
1551-6865  
入藏号:  
WOS:001011190000011  
Indexed Date:  
2023-07-27  
Institution Address:  
City Univ Hong Kong, Hong Kong, Peoples R China  
Institution Address:  
East China Normal Univ, Shanghai, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Shanghai, Peoples R China  
Institution:  
Institution:  
Institution:

### 902. Bi-directional skip connection feature pyramid network and sub-pixel

摘要: In existing state-of-the-art object detectors, feature pyramid networks (FPN) and multiscale feature fusion are still typically used. The traditional FPN fusion strategy is based on the top-down fusion of high-level semantic information. The top-down fusion method generally uses upsampling based on interpolation, which often results in jagged edges, mosaic distortion, and edge blurring. Moreover, in order to improve accuracy, the FPN-based fusion strategy must add multiple top-down components for fusion, which increases computational costs and leads to a poor balance between precision and speed. In this paper, we propose a novel fusion strategy based on a backbone network. We aim to design simple and efficient components for high-quality object detection. Our proposed strategy, bi-directional skip connection FPN (BiSCFPN), consists of three components: a bi-directional skip connection (BiSC), a selective dilated convolution module (SDCM), and sub-pixel convolution (SP). The BiSC aims to enhance semantic information between different feature layers in the backbone network and simultaneously uses the SDCM to improve the receptive fields of differently sized targets in the fusion stage. Finally, SP learns the relationship between the features of upsampling and downsampling images to effectively mitigate the problems caused by the traditional interpolation method. BiSCFPN achieves an average precision of 38.2% in tests with the Microsoft Common Objects in Context (MS COCO) test-dev dataset at a real-time speed of-50 FPS & eth;608 x 608 & THORN; using an Nvidia GeForce RTX 2080 Ti graphics card and significantly improves the balance between precision and speed.  
(c) 2021 Published by Elsevier B.V.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
被引频次合计:  
13  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000642403400001  
Indexed Date:  
2021-05-24  
Institution Address:  
Sichuan Univ, Coll Elect & Informat Engn, Chengdu 610065, Peoples R China  
Institution Address:  
Civil Aviat Adm China, Res Inst 2, Chengdu 610041, Peoples R China  
Institution:  
Institution:  
Civil Aviat Adm China

### 903. MAFNet: A deep multi-scale attentive fusion network for virtual

摘要: An essential step for 3D virtual surgical planning in orthognathic surgery is image segmentation to generate virtual 3D models for virtual osteotomy of the maxillofacial bone in CT images. However, this manual segmentation process is time-consuming and labor-intensive and requires expertise. Also, most conventional automatic segmentation methods can not directly segment the maxillofacial bone for virtual Le Fort I osteotomy. The purpose of this study was to automatically and robustly segment the maxillofacial bone for virtual Le Fort I osteotomy in CT images using a deep multi -scale attentive fusion network (MAFNet). MAFNet consisted of multi -scale encoders (MEs), an atrous feature fusion module (AFFM), a multi -scale spatial attention module (MSAM), and a weighted soft combo loss (SCL) with deep supervision. In performance comparisons, MAFNet outperformed popular segmentation networks by achieving 0.836, 0.951, and 0.929 in terms of F1 -score for the maxilla, mandible, and skull segmentation, respectively. In ablation studies, the proposed components and SCL improved segmentation performance by allowing the network to learn multiscale feature representations and anatomical contexts efficiently. MEs and AFFM captured multi -scale context features and mitigated the loss of spatial information. MSAM emphasized the boundary details, such as an osteotomy line between the maxilla and skull, and suppressed irrelevant background details, such as metal artifacts and soft tissues. SCL with deep supervision helped MAFNet alleviate the class imbalance problem and improve segmentation performance. Therefore, MAFNet outperformed popular segmentation networks and obtained robust segmentation results on challenging CT images including osteotomy lines, metal artifacts, and malocclusions in CT images.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:001243015500001  
Indexed Date:  
2024-06-16  
Institution Address:  
Seoul Natl Univ, Grad Sch Convergence Sci & Technol, Dept Appl Bioengn, Seoul 08826, South Korea  
Institution Address:  
Seoul Natl Univ, Grad Sch Convergence Sci & Technol, Dept Biomed Radiat Sci, Seoul 08826, South Korea  
Institution Address:  
LG CNS, AI Engn Team, Seoul, South Korea  
Institution Address:  
Hansung Univ, Dept Elect & Informat Engn, Seoul 02876, South Korea  
Institution Address:  
Seoul Natl Univ, Sch Dent, Dept Oral & Maxillofacial Radiol, Seoul 03080, South Korea  
Institution Address:  
Seoul Natl Univ, Dent Res Inst, Sch Dent, Seoul 03080, South Korea  
Institution:  
Institution:  
Institution:  
LG CNS  
Institution:  
Institution:  
Institution:

### 904. Re-parameterization Enhanced Dual-modal Realtime Object Detection Model

摘要: The objects captured by drones at high altitudes are generally small and have weak features,and they are greatly affec-ted by complex weather conditions.Object detection based on visible or infrared images often has high rates of missed detection and false detection.To address this problem,this paper proposes a dual-modal realtime object detection model DM-YOLO with reparameterization enhancement.Firstly,the visible and infrared images are effectively fused by channel concatenation,which makes efficient use of the complementary information in the dual-modal images at a very low cost.Secondly,a more efficient reparameterization module is proposed and a more powerful backbone network RepCSPDarkNet is constructed based on it,which effectively improves the feature extraction capability of the backbone network for dual-modal images.Then,a multi-level feature fusion module is proposed to enhance the multiscale feature representation of weak and small objects by fusing multi-scale feature information of weak and small objects with multi-receptive field dilated convolution and attention mechanism.Finally,the deep feature layer of the feature pyramid is removed,which reduces the model size while maintaining the detection accuracy.Experimental results on the large-scale dual-modal image dataset DroneVehicle show that,the detection accuracy of DM-YOLO is 2.45% higher than that of the baseline YOLOv5s,and is better than that of the YOLOv6 and YOLOv7 models.Furthermore,it effectively improves the accuracy and robustness of object detection under complex weather conditions,while achieving a detection speed of 82 frames per second,which can meet the requirements of realtime detection.  
摘要:  
无人机高空航拍的目标普遍尺寸小、特征弱,而且受复杂天候条件影响大,导致基于可见光或红外单模态图像的目标检测漏检、误检率较高。对此,提出了重参数化增强的双模态实时目标检测模型DM-YOLO。首先,采用通道拼接的方法融合可见光和红外图像,以极低的成本融合双模态图像的互补信息。其次,提出更加高效的重参数化模块并基于此构建了更加强大的骨干网RepCSPDarkNet,有效增强了骨干网对双模态图像的特征提取能力。然后,提出了多层次特征融合模块,通过多感受野卷积和注意力机制融合弱小目标的多尺度特征信息,增强了弱小目标的多尺度特征表示。最后,删除了对弱小目标检测基本不起作用的特征金字塔深层检测层,在检测精度保持不变的情况下,减小了模型规模。实验结果表明,在大规模的双模态图像数据集DroneVehicle上,DM-YOLO的检测精度比基准YOLOv5s高出2.45%,且优于规模相当的YOLOv6和YOLOv7模型,有效提高了复杂光照条件下目标检测的准确性和鲁棒性,同时检测速度达到82FPS,可满足实时检测的需求。  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-137X  
入藏号:  
CSCD:7820865  
Indexed Date:  
2024-12-28  
Institution Address:  
陆军工程大学指挥控制工程学院, 南京, 江苏 210007, 中国  
Institution Address:  
College of Command and Control Engineering,Army Engineering University of PLA, Nanjing, Jiangsu 210007, China  
Institution:  
陆军工程大学指挥控制工程学院  
Institution:  
College of Command and Control Engineering,Army Engineering University of PLA

### 905. RTLNet: Recursive Triple-Path Learning Network for Scene Parsing of

摘要: Scene parsing approaches have attracted extensive attention in recent years; although several methods have been developed for scene parsing, most include complex modules for both cross-modality fusion between RGB and depth images in the encoder and image scale level recovery in the decoder under label supervision for high inference accuracy. Cross-modality information in the encoder may be diluted when processed through the decoder, and the supervision results may not be reused effectively, which adversely affects scene parsing. To address these problems, we propose a recursive triple-path learning network (RTLNet) for cross-modality interactions in the decoder using global context and cross-modality fusion modules. The proposed modules fully use cross-modality information to reduce information loss. To enhance the robustness of RTLNet, we add a path to reuse the initial predictions from the decoder and introduce a ladder-shaped feature consistency module to further leverage multiscale features. Experiments are conducted with the proposed RTLNet and nine recent RGB-D indoor scene parsing methods on the NYUv2 and SUN-RGBD indoor scene datasets; the results show that the RTLNet outperforms the other methods.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1070-9908  
eISSN:  
1558-2361  
入藏号:  
WOS:000748371000037  
Indexed Date:  
2022-02-06  
Institution Address:  
Zhejiang Univ Sci & Technol, Sch Informat & Elect Engn, Hangzhou 310023, Peoples R China  
Institution Address:  
Zhejiang Univ, Coll Informat & Elect Engn, Hangzhou 310027, Peoples R China  
Institution:  
Institution:

### 906. Semantics and Contour Based Interactive Learning Network for Building

摘要: Building footprint extraction plays an important role in the analysis of remote-sensing images and has an extensive range of applications. Obtaining precise boundaries of buildings remains a challenge in existing building extraction methods. Some previous works have made notable efforts to address this concern. However, most of these methods require cumbersome and expensive postprocessing steps. Moreover, they ignored the correlation between building semantics and contours, which we believe is crucial for building footprint extraction. To mitigate this issue, this article presents an intuitive and effective framework that explores semantic and contour cues of buildings and fully excavates their correlation. Specifically, we construct an interactive dual-stream decoder. The intermediate connections within this decoder interactively transmit features between branches, contributing to learning correlations between semantics and contours. We propose the semantic collaboration module (SCM) to strengthen the connection between the two branches. To further boost performance, we build the multiscale semantic context fusion module (MSCF) to fuse semantic information from the higher and lower layers of the network, allowing the network to obtain superior feature representations. The experiment results on the WHU, INRIA, and Massachusetts building datasets demonstrate the superior performance of our method.  
Web of Science 核心合集中的 "被引频次":  
3  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001097583200008  
Indexed Date:  
2023-12-01  
Institution Address:  
Xidian Univ, Sch Artificial Intelligence, Xian 710071, Peoples R China  
Institution Address:  
Univ Leicester, Sch Comp & Math Sci, Leicester LE1 7RH, England  
Institution:  
Institution:

### 907. Fast and Accurate Visual Tracking with Group Convolution and Pixel-Level

摘要: Visual object trackers based on Siamese networks perform well in visual object tracking (VOT); however, degradation of the tracking accuracy occurs when the target has fast motion, large-scale changes, and occlusion. In this study, in order to solve this problem and enhance the inference speed of the tracker, fast and accurate visual tracking with a group convolution and pixel-level correlation based on a Siamese network is proposed. The algorithm incorporates multi-layer feature information on the basis of Siamese networks. We designed a multi-scale feature aggregated channel attention block (MCA) and a global-to-local-information-fused spatial attention block (GSA), which enhance the feature extraction capability of the network. The use of a pixel-level mutual correlation operation in the network to match the search region with the template region refines the bounding box and reduces background interference. Comparing our work with the latest algorithms, the precision and success rates on the UAV123, OTB100, LaSOT, and GOT10K datasets were improved, and our tracker was able to run at 40FPS, with a better performance in complex scenes such as those with occlusion, illumination changes, and fast-motion situations.  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001063701000001  
Indexed Date:  
2023-09-21  
Institution Address:  
Chinese Acad Sci, Changchun Inst Opt Fine Mech & Phys, Changchun 130033, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 908. AdaptoMixNet: detection of foreign objects on power transmission lines

摘要: With the expansion of power transmission line scale, the surrounding environment is complex and susceptible to foreign objects, severely threatening its safe operation. The current algorithm lacks stability and real-time performance in small target detection and severe weather conditions. Therefore, this paper proposes a method for detecting foreign objects on power transmission lines under severe weather conditions based on AdaptoMixNet. First, an Adaptive Fusion Module (AFM) is introduced, which improves the model's accuracy and adaptability through multi-scale feature extraction, fine-grained information preservation, and enhancing context information. Second, an Adaptive Feature Pyramid Module (AEFPM) is proposed, which enhances the focus on local details while preserving global information, improving the stability and robustness of feature representation. Finally, the Neuron Expansion Recursion Adaptive Filter (CARAFE) is designed, which enhances feature extraction, adaptive filtering, and recursive mechanisms, improving detection accuracy, robustness, and computational efficiency. Experimental results show that the method of this paper exhibits excellent performance in the detection of foreign objects on power transmission lines under complex backgrounds and harsh weather conditions.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001310907900002  
Indexed Date:  
2024-09-18  
Institution Address:  
Xian Polytech Univ, Xian 710048, Peoples R China  
Institution Address:  
Xian Key Lab Interconnected Sensing & Intelligent, Xian 710048, Peoples R China  
Institution Address:  
Xidian Univ, Xian 710126, Peoples R China  
Institution:  
Institution:  
Xian Key Lab Interconnected Sensing & Intelligent  
Institution:

### 909. A novel multi-attention, multi-scale 3D deep network for coronary artery

摘要: Automatic segmentation of coronary arteries provides vital assistance to enable accurate and efficient diagnosis and evaluation of coronary artery disease (CAD). However, the task of coronary artery segmentation (CAS) remains highly challenging due to the large-scale variations exhibited by coronary arteries, their complicated anatomical structures and morphologies, as well as the low contrast between vessels and their background. To comprehensively tackle these challenges, we propose a novel multi-attention, multi-scale 3D deep network for CAS, which we call CAS-Net. Specifically, we first propose an attention-guided feature fusion (AGFF) module to efficiently fuse adjacent hierarchical features in the encoding and decoding stages to capture more effectively latent semantic information. Then, we propose a scale-aware feature enhancement (SAFE) module, aiming to dynamically adjust the receptive fields to extract more expressive features effectively, thereby enhancing the feature representation capability of the network. Furthermore, we employ the multi-scale feature aggregation (MSFA) module to learn a more distinctive semantic representation for refining the vessel maps. In addition, considering that the limited training data annotated with a quality golden standard are also a significant factor restricting the development of CAS, we construct a new dataset containing 119 cases consisting of coronary computed tomographic angiography (CCTA) volumes and annotated coronary arteries. Extensive experiments on our self-collected dataset and three publicly available datasets demonstrate that the proposed method has good segmentation performance and generalization ability, outperforming multiple state-of-the-art algorithms on various metrics. Compared with U-Net3D, the proposed method significantly improves the Dice similarity coefficient (DSC) by at least 4% on each dataset, due to the synergistic effect among the three core modules, AGFF, SAFE, and MSFA. Our implementation is released at https://github.com/Cassie-CV/CAS-Net.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
27  
被引频次合计:  
27  
ISSN:  
1361-8415  
eISSN:  
1361-8423  
入藏号:  
WOS:000925338900001  
PubMed ID:  
36630869  
Indexed Date:  
2023-02-25  
Institution Address:  
Xi An Jiao Tong Univ, Affiliated Hosp 2, Inst Med Artificial Intelligence, Xian 710004, Peoples R China  
Institution Address:  
Shaanxi Normal Univ, Sch Comp Sci, Xian 710062, Peoples R China  
Institution:  
Institution:

### 910. Object Detection Based on Improved YOLOv7 for UAV Aerial Image

摘要: An improved YOLOv7 aerial image object detection algorithm is proposed to solve the problems of low detection accuracy caused by mesoscale changes, small targets and dense occlusion in UAV aerial images. Firstly, a weighted sampling module with joint dynamic convolution is designed to capture features from multiple dimensions and improve the feature extraction ability of the model. Secondly, add a shallow feature detection head to retain more detailed information and enhance the ability to utilize small target features. Then, a multi-scale feature aggregation module (C2-Res2Block) with residual structure is constructed in the feature fusion part to make the model fuse rich multi-scale information. Finally, the MPDIoU measure is used to replace the traditional IOU to calculate the boundary regression loss and improve the localization ability of the model to the densely occluding target. Experiments on UAV aerial photography data set VisDrone2019 show that the improved algorithm is 4.3 percentage points higher than the original model on mAP@0.5, 2.4 percentage points on mAP@0.5:0.95, the number of parameters is reduced by 6.81\*10~6, and the detection accuracy is higher than the current mainstream object detection algorithms. It effectively improves the detection accuracy of UAV aerial images, and obviously improves the false detection and missing detection of aerial objects.  
摘要:  
针对无人机航拍图像中尺度变化、小目标且密集遮挡情况导致检测精度降低的问题,提出一种改进YOLOv7的航拍图像目标检测算法。设计一种联合动态卷积的加权采样模块,从多个维度捕捉特征,提高模型特征提取能力;增加浅层特征检测头,保留更多的细节信息,增强对小目标特征的利用能力;在特征融合部分构建一种具有残差结构的多尺度特征聚合模块(C2-Res2Block),使模型融合丰富的多尺度信息。使用MPDIoU度量替换传统IOU计算边界回归损失,提高模型对密集遮挡目标的定位能力。通过在无人机航拍数据集VisDrone2019上进行实验表明,改进后的算法较原模型mAP@0.5提高了4.3个百分点,mAP@0.5:0.95提高了2.4个百分点,参数量减少了6.81\*10~6,与目前主流的目标检测算法相比也取得更高的检测精度,有效提高了对无人机航拍图像的检测精度,并明显改善对航拍目标的误检和漏检情况。  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7828716  
Indexed Date:  
2025-01-07  
Institution Address:  
辽宁工程技术大学软件学院, 葫芦岛, 辽宁 125105, 中国  
Institution Address:  
College of Software,Liaoning Technical University, Huludao, Liaoning 125105, China  
Institution:  
辽宁工程技术大学软件学院  
Institution:  
College of Software,Liaoning Technical University

### 911. Attention based lightweight asymmetric network for real-time semantic

摘要: Real-time semantic segmentation is one of the important tasks in the field of computer vision, which is widely used in the fields of autonomous driving and medical imaging. Existing lightweight networks usually improve inference speed at the sacrifice of segmentation accuracy. How to achieve a balance between accuracy and speed is still a challenging problem for real-time semantic segmentation. In this paper, we propose an attention based lightweight asymmetric network (ALANet) to address this problem. Specifically, in the encoder, a channel-wise attention based depth-wise asymmetric block (CADAB) is designed to extract sufficient features, which has a small number of parameters. In the decoder, a spatial attention based pyramid pooling (SAPP) module is presented to aggregate multi-scale context information by using a few convolutions and poolings; and a pixel-wise attention based multi-scale feature fusion (PAMFF) module is developed to fuse features from different scales and generate pixel-wise attention for improving image restoration. Our ALANet has only 1.32M parameters. Experimental results on the Cityscapes and CamVid datasets show that ALANet obtains the segmentation accuracy (mIoU) of 74.4% and 69.5% and the inference speed of 115.6FPS and 113.2FPS, respectively. These results demonstrate that ALANet achieves a good balance between accuracy and speed.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001149729700001  
Indexed Date:  
2024-02-05  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Artificial Intelligence, Sch Future Technol, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Comp Sci, Nanjing 210044, Peoples R China  
Institution:  
Institution:

### 912. Enhanced semantic feature pyramid network for small object detection

摘要: Feature-pyramid network-based models, which progressively fuse multi-scale features, have been proven highly effective in object detection. However, these models often learn multi-scale features with ambiguous boundaries, due to small objects with only a few pixels that easily lose information during top-down propagation, which makes multi-scale feature representation less effective. In this work, we propose an efficient Enhanced Semantic Feature Pyramid Network(ES-FPN), which combines semantic information at high-level with contextual information at low-level to improve multi-scale feature learning in small object detection. Specifically, the proposed network first exploits the rich semantic information in lateral connections that enables the features to be more semantic. Then, it excavates the lost information in high-level/low-res feature maps with rich contextual information in low-level/high-res. In this way, the high-level layers suffer the reduced loss of important contextual information during the progressive feature fusion that avoids object disappearance, which is useful to utilize rich semantic information in high-level. Finally, ES-FPN fuses the distributed features of each layer stage-by-stage and the final features are more semantically and better for localizing the object. Extensive experimental results over three widely used object detection benchmarks(MS COCO, VOC and Cityscapes) demonstrate that our network can accurately locate fairly complete objects with clear boundaries and outperforms previous feature pyramid-based methods.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
17  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
17  
ISSN:  
0923-5965  
eISSN:  
1879-2677  
入藏号:  
WOS:000995332800001  
Indexed Date:  
2023-06-11  
Institution Address:  
Zhejiang Normal Univ, Sch Math & Comp Sci, Jinhua 321004, Peoples R China  
Institution Address:  
Jiaxing Univ, Sch Informat Sci & Engn, Jiaxing 314001, Peoples R China  
Institution:  
Institution:

### 913. Hand-raising gesture detection in classroom with spatial context

摘要: Behavior is a manner to express one's intent, and it is a hotspot and tough problem under discussion in the field of computer vision. This research mainly focuses on behavior in educational scenes, with hand-raising being the most common in classrooms. The behavior of students raising their hands can be used to assess a student's involvement, teaching quality, classroom atmosphere and so on. To detect student's hand-raising in real classroom, we build a hand-raising dataset, and improve the baseline by boosting feature expression. First, we analyze the drawback of multi-scale feature fusion part in the baseline. Considering that the information loss is due to the reduction in the number of channels during feature pyramid construction, Spatial Context Augmentation (SCA) is applied to mitigate feature map information loss at the highest level. Then, when raising the hand, there is a spatial relationship between the head and the hand. Based on this breakthrough, we propose Multi-Branch Dilated Convolution (MBDC) to enlarge the receptive field and reduce false detection. Extensive experimental results on the hand-raising dataset show that our strategy outperforms a number of state-of-the-art approaches. In addition, our algorithm is also tested on PASCAL VOC, Cityscapes dataset. The results outperform other algorithms on both datasets. This also demonstrates the universality of our algorithm. (c) 2022 Published by Elsevier Ltd.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
8  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
ISSN:  
0097-8493  
eISSN:  
1873-7684  
入藏号:  
WOS:000924865900001  
Indexed Date:  
2023-02-25  
Institution Address:  
Sun Yat Sen Univ, Sch Elect & Commun Engn, Shenzhen 518107, Peoples R China  
Institution:

### 914. Road small target detection based on improved YOLO v5algorithm

摘要: In order to solve the problems that small targets on traffic roads faces including detection difficulty,low precision,detection failures,a multi-scale feature fusion target detection improvement algorithm based on the YOLO v5(you only look once v5)algorithm is proposed.Firstly,the small target detection head is added for adapting to the small target size and alleviating the missed detection.Then,deformable convolutional networks V2(DCN V2)is introduced to improve the model's learning ability for small targets in motion.The context augmentation module(CAM)is introduced to improve the recognition ability of small targets at a long distance.The replacement loss function is used to improve the bounding box's localization accuracy,and the spatial pyramid pooling and context spatial pyramid convolution\_group(SPPCSPC\_group) module is also used to improve the sensory field and feature expression ability of the network.The experiment results show that the proposed algorithm achieves an average accuracy of 95.2%in the category of small targets in the KITTI dataset,compared with the original YOLO v5algorithm,the overall average accuracy is improved by 2.7%.For the detection of small targets,the average accuracy is improved by 3.1% with a better detection effect,which proves the effectiveness of the proposed algorithm for the detection of small targets on roads.  
摘要:  
为解决交通道路小目标检测难度大、精度低,容易出现错检漏检的问题,提出一种基于YOLO v5 (you only look once v5)算法的多尺度特征融合目标检测改进算法。首先,增加小目标检测头用于适应小目标尺寸,缓解漏检情况。然后,引入可变形卷积网络v2(deformable convolutional networks V2,DCN V2)提高模型对运动中小目标的学习能力;同时,增加上下文增强模块,提升对远距离小目标的识别能力。最后,在替换损失函数、提高边界框定位精度的同时,使用空间金字塔池化和上下文空间金字塔卷积分组模块,提高网络的感受野和特征表达能力。实验结果表明,所提算法在KITTI数据集小目标类别上平均识别精度达到了95.2%,相较于原始YOLO v5,算法总体平均识别精度提升了2.7%,对小目标的检测效果更佳,平均识别精度提升了3.1%,证明所提算法在道路小目标检测方面的有效性。  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1001-506X  
入藏号:  
CSCD:7818269  
Indexed Date:  
2024-12-21  
Institution Address:  
大连交通大学软件学院, 大连, 辽宁 116028, 中国  
Institution Address:  
School of Software,Dalian Jiaotong University, Dalian, Liaoning 116028, China  
Institution:  
大连交通大学软件学院  
Institution:  
School of Software,Dalian Jiaotong University

### 915. Detection algorithm for dense small objects in high altitude image

摘要: The high altitude images captured by drones often contain small-sized objects, which pose several challenges including significant scale variations and dense clustering. To tackle these challenges in small object detection, this paper presents EM-YOLO (Efficient Switching Network and Multi-Scale Feature Fusion Reinforced YOLO for Small Object Detection), an enhanced algorithm based on the YOLOv5s architecture. Firstly, to enhance the network's adaptability to varying object feature scales, the EM-YOLO redesigns the feature extraction backbone network by using different dilation rates in convolutions within the backbone. This flexible structure allows for improved extraction of object features, enhancing the network's feature extraction capability. Secondly, additional detection layer for small objects and combines it with the dynamic detection head. This combination addresses the issue of small object information loss due to multiple down-sampling operations on the feature maps. Finally, Focal-EIoU Loss is employed to enhance regression accuracy in EM-YOLO. Extensive experiments were conducted on the VisDrone2019-DET dataset. The results of these experiments demonstrate that the improved algorithm achieved a significant increase in performance. Specifically, the algorithm achieves 43.5% and 25.1% on the mAP0.5 and mAP0.5 : 0.95 metrics, which are 13.9% and 8.9% higher than YOLOv5s algorithm, respectively. Furthermore, comparative experiments with other mainstream algorithms indicate that EM-YOLO outperforms them in the task of small object detection.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
1051-2004  
eISSN:  
1095-4333  
入藏号:  
WOS:001165299600001  
Indexed Date:  
2024-03-04  
Institution Address:  
Yunnan Normal Univ, Sch Informat, Kunming, Peoples R China  
Institution Address:  
Engn Res Ctr Comp Vis & Intelligent Control Techno, Dept Educ Yunnan Prov, Kunming, Peoples R China  
Institution:  
Institution:  
Engn Res Ctr Comp Vis & Intelligent Control Techno

### 916. Semantic-agnostic progressive subtractive network for image manipulation

摘要: In this paper, we propose a new detection and localization framework capable of detecting suspicious forgeries using the Semantic-Agnostic Progressive Subtractive Network (SAPS-Net). Our approach is based on the key observation that fluctuations in image content severely interfere with the capture of general manipulations by existing convolutional architecture. Distinct from the aggregation attention employed by traditional methods, we design the Semantic-Agnostic Manipulation Attention (SAMA) based on subtractive operation for mitigating the effect of rich image semantics on manipulation extraction. Initially, the Multi-Scale feature Iterative Fusion Block (MSIFB) and Multi-Kernel feature Fusion Residual Block (MKFRB) are designed to iteratively crawl potential semantic associations of different hierarchical feature mappings. Then, we further devise the subtractive operation to effectively remove the semantic associations as distractors and promote the network to adaptively learn general forgery. Notably, these semantic associations based on image content may be fundamentally different from the manipulation traces that alter the internal patterns of images. By progressively utilizing SAMAs, the network remains robust to image content manipulation with rich semantics. Extensive experiments on six challenge datasets show that our approach has more than 3.03% pixel-level AUC gains and 3.70% imagelevel AUC gains in cross-dataset scenarios compared to state-of-the-art methods, especially on the IMD20 dataset (pixel-level AUC: 0.859) and Wild dataset (pixel-level AUC: 0.821) with realistic scenarios.(c) 2023 Elsevier B.V. All rights reserved.  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000999586300001  
Indexed Date:  
2023-06-15  
Institution Address:  
Jilin Univ, Coll Comp Sci & Technol, Changchun 130012, Jilin, Peoples R China  
Institution Address:  
Jilin Univ, Key Lab Symbol Computat & Knowledge Engn, Minist Educ, Changchun 130012, Jilin, Peoples R China  
Institution Address:  
Hulunbuir Univ, Coll Comp, Hulunbuir 021008, Inner Mongolia, Peoples R China  
Institution:  
Institution:  
Institution:

### 917. Named Entity Recognition Based on Multi-scale Attention

摘要: The accuracy of named entity recognition(NER) task will promote the research of multiple downstream tasks in natural language field. Due to a large number of nested semantics in text, named entities are recognized difficultly. Recognizing nested semantics becomes a difficulty in natural language processing. Previous studies have single scale of extracting feature and under-utilization of the boundary information. They ignore many details under different scales and then lead to the situation of entity recognition error or omission. Aiming at the above problems, a multi- scale attention method for named entity recognition(MSA-NER) is proposed. Firstly, the BERT model is used to obtain representation vector containing context information, and then the BiLSTM network is used to strengthen the context representation of text. Secondly, the representation vectors are enumerated and concatenated to form span information matrix. The direction information is fused to obtain richer interactive information. Thirdly, multi-head attention is used to construct multiple subspaces. Two-dimensional convolution is used to optionally aggregate text information at different scales in each subspace, so as to implement multi-scale feature fusion in each attention layer. Finally, the fused matrix is used for span classification to identify named entities. Experimental results show that the F1 score of the proposed method reaches 81.7% and 86.8% on GENIA and ACE2005 English datasets, respectively. The proposed method demonstrates better recognition performance compared with existing mainstream models.  
摘要:  
命名实体识别(NER)任务的准确性将促进自然语言领域中诸多下游任务的研究。由于文本中存在大量嵌套语义,导致命名实体识别困难,成为自然语言处理中的难点。以往研究提取特征尺度单一,边界信息利用不够充分,忽略了不同尺度下的许多细节信息,从而造成实体识别错误或遗漏的情况。针对上述问题,提出一种多尺度注意力的命名实体识别方法(MSA-NER)。首先,利用BERT模型得到包含上下文信息的表示向量,并通过BiLSTM网络加强文本的上下文表示。其次,将表示向量进行枚举拼接形成跨度信息矩阵,并融合方向信息获得更丰富的交互信息。然后,利用多头注意力构建多个子空间,通过二维卷积在每个子空间下可选地聚合不同尺度的文本信息,在每个注意力层同时进行多尺度的特征融合。最后,将融合的矩阵进行跨度分类以识别命名实体。实验表明,该方法在GENIA和ACE2005英文数据集上F1分别达到81.7%和86.8%,与现有主流模型相比有更好的识别效果。  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1673-9418  
入藏号:  
CSCD:7648953  
Indexed Date:  
2024-04-12  
Institution Address:  
贵州财经大学信息学院;;贵州大学计算机科学与技术学院;;公共大数据国家重点实验室, ;;;;公共大数据国家重点实验室, 贵阳;;贵阳;;贵阳, ;;;; 550025;;550025;;550025  
Institution Address:  
贵州大学计算机科学与技术学院;;公共大数据国家重点实验室, ;;公共大数据国家重点实验室, 贵阳;;贵阳, ;; 550025;;550025  
Institution Address:  
School of Information, Guizhou University of Finance and Economics;;College of Computer Science and Technology, Guizhou University;;State Key Laboratory of Public Big Data, ;;;;State Key Laboratory of Public Big Data, Guiyang;;Guiyang;;Guiyang, ;;;; 550025;;550025;;550025  
Institution Address:  
College of Computer Science and Technology, Guizhou University;;State Key Laboratory of Public Big Data, ;;State Key Laboratory of Public Big Data, Guiyang;;Guiyang, ;; 550025;;550025  
Institution:  
贵州财经大学信息学院;;贵州大学计算机科学与技术学院;;公共大数据国家重点实验室  
Institution:  
贵州大学计算机科学与技术学院;;公共大数据国家重点实验室  
Institution:  
School of Information, Guizhou University of Finance and Economics;;College of Computer Science and Technology, Guizhou University;;State Key Laboratory of Public Big Data  
Institution:  
College of Computer Science and Technology, Guizhou University;;State Key Laboratory of Public Big Data

### 918. Concrete dam apparent crack characteristic extraction based on improved

摘要: To address the problems of difficulty and poor effect of existing crack detection algorithm caused by complex environment of concrete dams, a characteristic extraction method for the cracks of concrete dams based on the improved DeepLabv3 + model was proposed. This method replaced the original backbone network with lightweight network to extract image features, which reduced the complexity of the model. The atrous spatial pyramid pooling module was expanded to widen the encoder receptive field. A multi-scale feature fusion strategy was adopted to improve the utilization of edge information. Moreover, the loss function of the model was optimized to overcome the pixel imbalance. The effectiveness and superiority of the proposed method were verified and evaluated using the self-made concrete dam apparent crack image dataset. The experimental results demonstrate that the proposed network can accurately retrieval characteristics of concrete dam apparent cracks under complex background. The intersection over union and pixel accuracy of segmented crack images are 72.85% and 85.36%, respectively. Compared with other classical image segmentation models, the proposed network has a significantly more prominent crack detection effect. It can provide an effective technical support for long-term concrete dam apparent crack monitoring.  
摘要:  
为了解决混凝土坝环境复杂造成现有算法裂缝检测难度大、效果差的问题,提出了一种改进Deep-Labv3 +模型的混凝土坝裂缝特征提取方法.该方法以轻量型网络替换原始骨干网络提取图像特征,降低模型复杂度;扩充空洞空间金字塔池化模块,提升编码器感受野;采用多尺度特征融合策略,提高边缘信息利用率;优化模型损失函数,克服像素不均衡的困难.采用自制混凝土坝表观裂缝图像数据集对提出方法的有效性和优越性进行了验证与评估,结果表明:构建的改进网络能准确地实现复杂背景下混凝土坝表观裂缝特征的提取,分割裂缝图像的交并比与像素精度分别为72.85%与85.36%,裂缝分割效果也明显优于其他方法,可为长期混凝土坝面裂缝监测提供有效的技术手段.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1001-0505  
入藏号:  
CSCD:7785427  
Indexed Date:  
2024-10-25  
Institution Address:  
沈阳建筑大学土木工程学院, 沈阳, 辽宁 110168, 中国  
Institution Address:  
安徽水安建设集团股份有限公司, 合肥, 安徽 230601, 中国  
Institution Address:  
大连理工大学水利工程学院;;河海大学水利水电学院, ;;, 大连;;南京, ;; 116023;;210098  
Institution Address:  
School of Civil Engineering,Shenyang Jianzhu University, Shenyang, Liaoning 110168, China  
Institution Address:  
Anhui Shui'an Construction Group Co.,Ltd., Hefei, Anhui 230601, China  
Institution Address:  
School of Hydraulic Engineering,Dalian University of Technology;;College of Water Conservancy and Hydropower Engineering,Hohai University, ;;, Dalian;;Nanjing, ;; 116023;;210098  
Institution:  
沈阳建筑大学土木工程学院  
Institution:  
安徽水安建设集团股份有限公司  
Institution:  
大连理工大学水利工程学院;;河海大学水利水电学院  
Institution:  
School of Civil Engineering,Shenyang Jianzhu University  
Institution:  
Anhui Shui'an Construction Group Co.,Ltd.  
Institution:  
School of Hydraulic Engineering,Dalian University of Technology;;College of Water Conservancy and Hydropower Engineering,Hohai University

### 919. A multi-scale feature extraction and fusion-based model for retinal

摘要: In response to the challenge of low accuracy in retinal vessel segmentation attributed to the minute nature of the vessels, this paper proposes a retinal vessel segmentation model based on an improved U-Net, which combines multi-scale feature extraction and fusion techniques. An improved dilated residual module was first used to replace the original convolutional layer of U-Net, and this module, coupled with a dual attention mechanism and diverse expansion rates, facilitates the extraction of multi-scale vascular features. Moreover, an adaptive feature fusion module was added at the skip connections of the model to improve vessel connectivity. To further optimize network training, a hybrid loss function is employed to mitigate the class imbalance between vessels and the background. Experimental results on the DRIVE dataset and CHASE\_DB1 dataset show that the proposed model has an accuracy of 96.27% and 96.96%, sensitivity of 81.32% and 82.59%, and AUC of 98.34% and 98.70%, respectively, demonstrating superior segmentation performance.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0140-0118  
eISSN:  
1741-0444  
入藏号:  
WOS:001337228600001  
PubMed ID:  
39432221  
Indexed Date:  
2024-10-27  
Institution Address:  
Southwest Univ Sci & Technol, Sch Informat Engn, Mianyang 621000, Peoples R China  
Institution Address:  
Robot Technol Used Special Environm Key Lab Sichua, Mianyang 621000, Peoples R China  
Institution:  
Institution:  
Robot Technol Used Special Environm Key Lab Sichua

### 920. Traffic Sign Recognition Model with Long-Tail Distribution Based on

摘要: Accurate recognition of traffic signs plays an important role in the field of intelligent driving. Traffic sign training datasets with long-tail distribution increase the difficulty of traffic sign recognition. A traffic sign recognition model with long-tail distribution based on YOLOX-Tiny was proposed to improve the poor performance of the model trained on long-tail distribution datasets. A long-tail traffic sign dataset was created based on the TT100K\_2021(tsinghua-tencent 100K 2021) dataset. YOLOX-Tiny was chosen as the underlying model by considering picture numbers in datasets, sample distribution, and model size. Equalization loss v2(EQL v2) was used as classification loss to balance the head and tail of the classifier, and focal loss(FL) was used as target confidence loss to enhance the model's prediction of target confidence. In order to solve the backpropagation conflicts of feature graphs at different levels on the traditional feature pyramid, enhance the feature reorganization effect, and highlight target feature, up-sampling operator CARAFE, coordinate attention(CA), and CARAFE + adaptively spatial feature fusion modules(CAR-ASFF) were introduced to the neck bidirectional pyramid. The research results show that the improved YOLOX-Tiny model achieves 43.67% and 29.98% respectively in the long-tail traffic sign datasets, namely mAP\_(50) and mAP\_(50):95. The improved model has higher detection accuracy than other target detection models.  
摘要:  
在智能驾驶领域,准确识别交通标志对行车安全具有重要意义,交通标志训练集往往服从长尾分布,这为交通标志识别带来极大难度。针对于长尾分布数据集训练出的模型在尾类上表现差的现象,提出一种基于YOLOX-Tiny的长尾分布交通标志识别模型。在TT100K\_2021(tsinghua-tencent 100K 2021)数据集基础上制作交通标志长尾数据集;从制作数据集图片数量、样本分布以及模型大小出发,选择YOLOX-Tiny作为基础模型;采用EQL v2(equalization loss v2)和FL(focal loss)作为分类损失和目标置信度损失,平衡分类器头尾差距,增强模型对目标置信度的预测;在颈部双向金字塔中引入上采样算子CARAFE、坐标注意力机制(coordinate attention,CA)和CAR-ASFF模块(CARAFE + adaptively spatial feature fusion),解决传统特征金字塔上不同层级特征图的反向传播冲突问题,提升特征重组效果,突出目标特征。研究结果表明:改进的YOLOXTiny模型在制作的长尾交通标志数据集上mAP\_(50)和mAP\_(50:95)分别达到了43.67%和29.98%,改进模型相比较其他几种目标检测模型具有更高的检测精度。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1004-731X  
入藏号:  
CSCD:7843444  
Indexed Date:  
2025-01-25  
Institution Address:  
湖北大学, 武汉, 湖北 430000, 中国  
Institution Address:  
中国科学院自动化研究所, 北京 100190, 中国  
Institution Address:  
北京中科慧眼科技有限公司, 北京 100190, 中国  
Institution Address:  
Hubei University, Wuhan, Hubei 430000, China  
Institution Address:  
Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China  
Institution Address:  
Beijing Smarter Eye Technology Company, Beijing 100190, China  
Institution:  
湖北大学  
Institution:  
中国科学院自动化研究所  
Institution:  
北京中科慧眼科技有限公司  
Institution:  
Hubei University  
Institution:  
Institute of Automation, Chinese Academy of Sciences  
Institution:  
Beijing Smarter Eye Technology Company

### 921. A multi-task network for occluded meter reading with synthetic data

摘要: The efficient pointer meter reading methods have been proposed based on machine vision to replace timeconsuming manual inspections for the industrial monitoring. However, the interference factors, such as rain or dirt, can occlude meter, which poses obstacles in the recognition and labeling of pointer and scales. To solve these problems, we propose a multi-task network with pointer and main scale detection (PMSD-Net) for the occluded meter reading with synthetic data generation technology. Specifically, dense parallel dilated convolution block is proposed for correlating the pointer and main scale features with large receptive field. Multi-scale feature fusion is designed to purify noisy features for the detailed information extraction. The relation reconstruction mechanism is designed to reconstruct the feature relation under severe occlusion. Moreover, the keypoint detection branch is designed to detect meter center and pointer tip according to the segmented pointer, which can identify changeable position of the segmented pointer tip to determine the pointer orientation. Finally, the synthetic data generation technology is developed to generate massive labeled data with simulated interference factors in the meter for the training, which enhances the generalization ability of PMSD-Net in various occlusion scenes. Experimental results indicate that PMSD-Net can segment more accurate regions of pointer and main scale and detect the changeable position of pointer tip for occluded meters, thereby improving the accuracy in reading occluded meters.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1474-0346  
eISSN:  
1873-5320  
入藏号:  
WOS:001392909800001  
Indexed Date:  
2025-01-15  
Institution Address:  
Fuzhou Univ, Coll Elect Engn & Automat, Fuzhou 350108, Fujian, Peoples R China  
Institution Address:  
State Grid Fujian Elect Power Co Ltd, Fuzhou Power Supply Co, Fuzhou 350007, Fujian, Peoples R China  
Institution Address:  
Yanshan Univ, Sch Elect Engn, Qinhuangdao 066004, Hebei, Peoples R China  
Institution:  
Institution:  
Institution:

### 922. Driver fatigue state detection method based on multi-feature fusion

摘要: The improved YOLOv5 object detection algorithm was used to detect the facial region of the driver and a multi-feature fusion fatigue state detection method was established aiming at the problem that existing fatigue state detection method cannot be applied to drivers under the epidemic prevention and control. The image tag data including the situation of wearing a mask and the situation without wearing a mask were established according to the characteristics of bus driving. The detection accuracy of eyes, mouth and face regions was improved by increasing the feature sampling times of YOLOv5 model. The BiFPN network structure was used to retain multi-scale feature information, which makes the prediction network more sensitive to targets of different sizes and improves the detection ability of the overall model. A parameter compensation mechanism was proposed combined with face keypoint algorithm in order to improve the accuracy of blink and yawn frame number. A variety of fatigue parameters were fused and normalized to conduct fatigue classification. The results of the public dataset NTHU and the self-made dataset show that the proposed method can recognize the blink and yawn of drivers both with and without masks, and can accurately judge the fatigue state of drivers.  
摘要:  
针对现有疲劳状态检测方法无法适用于疫情防控下的驾驶员,利用改进后的YOLOv5目标检测算法,对驾驶员的面部区域进行检测,建立多特征融合的疲劳状态检测方法.针对公交驾驶特性,建立包含佩戴口罩和未佩戴口罩情况的图像标签数据.通过增加YOLOv5模型的特征采样次数,提高眼、嘴、面部区域的检测精度.利用BiFPN网络结构保留多尺度的特征信息,使得预测网络对不同大小的目标更敏感,提升整体模型的检测能力.结合人脸关键点算法提出参数补偿机制,提高眨眼、打哈欠帧数的准确率.将多种疲劳参数融合归一化处理,开展疲劳等级划分.公开数据集NTHU和自制数据集的验证结果表明,该方法对佩戴口罩和未佩戴口罩情况均可以进行眨眼、打哈欠识别,可以准确地判断驾驶员的疲劳状态.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1008-973X  
入藏号:  
CSCD:7518811  
Indexed Date:  
2023-11-05  
Institution Address:  
浙江工业大学智能交通系统联合研究所, 杭州, 浙江 310014, 中国  
Institution Address:  
杭州金通科技集团股份有限公司, 杭州, 浙江 310014, 中国  
Institution Address:  
ITS Joint Research Institute, Zhejiang University of Technology, Hangzhou, Zhejiang 310014, China  
Institution Address:  
Hangzhou Jintong Technology Group Limited Company, Hangzhou, Zhejiang 310014, China  
Institution:  
浙江工业大学智能交通系统联合研究所  
Institution:  
杭州金通科技集团股份有限公司  
Institution:  
ITS Joint Research Institute, Zhejiang University of Technology  
Institution:  
Hangzhou Jintong Technology Group Limited Company

### 923. Micro-Expression Spotting Based on a Short-Duration Prior and

摘要: When micro-expressions are mixed with normal or macro-expressions, it becomes increasingly challenging to spot them in long videos. Aiming at the specific time prior of micro-expressions (MEs), an ME spotting network called AEM-Net (adaptive enhanced ME detection network) is proposed. This paper is an extension of the conference paper presented at the Chinese Conference on Biometric Recognition (CCBR). The network improves spotting performance in the following five aspects. Firstly, a multi-stage channel feature extraction module is constructed to extract the features at different depths. Then, an attention spatial-temporal module is leveraged to obtain salient and discriminative micro-expression segments while suppressing the generation of excessively long or short suggestions. Thirdly, a ME-NMS (non-maximum suppression) network is developed to reduce redundancy and decision errors. Fourthly, a multi-scale feature fusion module is introduced to fuse up-sampling features of high-level maps and fine-grained information, which obtains meaningful information on feature distribution and contributes to a good representation of MEs. Finally, two spotting mechanisms named anchor-based and anchor free were integrated to get final spotting. Extensive experiments were conducted on prevalent CAS(ME)(2) and the SAMM-Long ME databases to evaluate the spotting performance. The results show that the AEM-Net achieves competitive performance, outperforming other state-of-the-art methods.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2079-9292  
入藏号:  
WOS:000919604200001  
Indexed Date:  
2023-02-12  
Institution Address:  
Jiangxi Sci & Technol Normal Univ, Key Lab Opt Elect & Commun, Nanchang 330031, Peoples R China  
Institution:

### 924. Unsupervised Video Person Re-identification Based on Multiple Kernel

摘要: Person re-identification aims to identify specific individuals across surveillance cameras, overcoming challenges such as pose variations, occlusions, and background noise that often lead to insufficient feature extraction. This paper proposes a novel unsupervised video-based person reidentification method that utilizes multi-kernel dilated convolution to provide a more comprehensive and accurate representation of individual differences and features. Initially, we employ a pre-trained ResNet50 as an encoder. To further enhance the encoder's feature extraction capability, we introduce a multiple kernel dilated convolution module. Enlarging the receptive field of convolutional kernels allows the network to more effectively capture both local and global feature information, offering a more comprehensive depiction of a person's appearance features. Subsequently, a decoder is employed to restore high-level semantic information to a more fundamental feature representation, thereby strengthening feature representation and improving system performance under complex imaging conditions. Finally, a multi-scale feature fusion module is introduced in the decoder output to merge features from adjacent layers, reducing semantic gaps between different feature channel layers and generating more robust feature representations. Offline experiments are conducted on three mainstream datasets, and results show that the proposed method achieves significant improvements in both accuracy and robustness.  
摘要:  
行人重识别旨在跨监控摄像头下检索出特定的行人目标。由于存在姿态变化、物体遮挡和背景干扰的不同成像条件等问题,导致行人特征提取不充分。本文提出一种利用多核扩展卷积的无监督视频行人重识别方法,使得提取到的行人特征能够更全面、更准确地表达个体差异和特征信息。首先,采用预训练的ResNet50作为编码器,为了进一步提升编码器的特征提取能力,引入了多核扩展卷积模块,通过增加卷积核的感受野,使得网络能够更有效地捕获到局部和全局的特征信息,从而更全面地描述行人的外貌特征;其次,通过解码器将高级语义信息还原为更为底层的特征表示,从而增强特征表示,提高系统在复杂成像条件下的性能;最后,在解码器的输出中引入多尺度特征融合模块融合相邻层中的特征,进一步减少不同特征通道层之间的语义差距,以产生更鲁棒的特征表示。在3个主流数据集上进行离线实验,结果表明该方法在准确性和鲁棒性上均取得了显著的改进。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1004-9037  
入藏号:  
CSCD:7822902  
Indexed Date:  
2024-12-28  
Institution Address:  
兰州理工大学电气工程与信息工程学院;;甘肃省工业过程先进控制重点实验室, ;;甘肃省工业过程先进控制重点实验室, 兰州;;兰州, ;; 730050;;730050  
Institution Address:  
西北民族大学数学与计算机科学学院, 兰州, 甘肃 730030, 中国  
Institution Address:  
School of Electrical Engineering and Information Engineering,Lanzhou University of Technology;;Key Laboratory of Gansu Advanced Control for Industrial Processes, ;;Key Laboratory of Gansu Advanced Control for Industrial Processes, Lanzhou;;Lanzhou, ;; 730050;;730050  
Institution Address:  
College of Mathematics and Computer Science,Northwest Minzu University, Lanzhou, Gansu 730030, China  
Institution:  
兰州理工大学电气工程与信息工程学院;;甘肃省工业过程先进控制重点实验室  
Institution:  
西北民族大学数学与计算机科学学院  
Institution:  
School of Electrical Engineering and Information Engineering,Lanzhou University of Technology;;Key Laboratory of Gansu Advanced Control for Industrial Processes  
Institution:  
College of Mathematics and Computer Science,Northwest Minzu University

### 925. Semi-Supervised Image Deraining Using Knowledge Distillation

摘要: Image deraining has achieved considerable progress based on supervised learning with synthetic training pairs, but is usually limited in handling real-world rainy images. Although semi-supervised methods are suggested to exploit real-world rainy images when training deep deraining models, their performances are still notably inferior. To address this crucial issue, this work proposes a semi-supervised image deraining network with knowledge distillation (SSID-KD) for better exploiting real-world rainy images. In particular, the consistency of feature distribution of rain streaks extracted from synthetic and real-world rainy images is enforced by adopting knowledge distillation. Moreover, as for the backbone in SSID-KD, we propose the multi-scale feature fusion module and the pyramid fusion module to better extract deep features of rainy images. SSID-KD can relieve the problem of over-deraining or under-deraining for real-world rainy images, while it can keep comparable performance with supervised deraining methods on several benchmark datasets. Extensive experiments on both synthetic and real-world rainy images have validated that our SSID-KD not only can achieve better deraining results than existing semi-supervised deraining methods but also are quantitatively comparable with state-of-the-art supervised deraining methods. Benefiting from the well exploration of real-world rainy images, our SSID-KD can obtain more visually plausible deraining results. The source code and trained models are publicly available at https://github.com/cuiyixin555/SSID-KD.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
37  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
5  
被引频次合计:  
42  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:000936985600020  
Indexed Date:  
2023-03-10  
Institution Address:  
Tianjin Univ, Coll Intelligence & Comp, Tianjin 300350, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Dept Comp, Kowloon, Hong Kong, Peoples R China  
Institution Address:  
Harbin Inst Technol, Sch Comp Sci & Technol, Harbin 150001, Peoples R China  
Institution Address:  
Alibaba Cloud, Hangzhou 310052, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 926. SoybeanTracer: An In-Field scene property-based framework for

摘要: Soybean is a crucial plant-based protein and vegetable oil source for the global population and a significant cereal-oil dual-purpose crop. Extracting soybean canopy coverage, subsequently implementing evaluation in a low-cost, high-throughput and accurate manner, is highly valuable for soybean breeding and yield increase. Although the utilization of RGB remote sensing images for soybean analysis offers distinct advantages in terms of its cost effectiveness, there are still issues to be addressed. Specifically, the in-field scene properties of soybean canopy data have rarely been developed and there is a gap in automatic evaluation of soybean canopy coverage. For purpose of tackling these pending issues, this study proposed SoybeanTracer, a novel framework designed to achieve accurate soybean canopy coverage extraction and enable specialized evaluation with low cost and high throughput. The proposed framework comprised two components: Soybean Canopy Segmentation Network (SCSNet) and an agronomic soybean coverage evaluation system. Specifically, SCSNet employed a feature extraction network integrating the homogeneity within plot regions and soybean distribution characteristics. To further address the segmentation challenge of soybean plants with sparse growth posed by genotypes, a soybean multiscale feature fusion module was designed. As a result, SCSNet achieved a mean intersection over union (mIoU) and mean pixel accuracy (mPA) of 89.39% and 94.42%, respectively, on soybean canopy segmentation. Therefore, accurate and low-cost extraction of soybean canopy coverage was achieved through the use of RGB remote sensing data. Furthermore, the soybean canopy coverage evaluation system proposed in this study encompassed five levels, characterized by wider spacing at the two ends and narrower spacing in the middle. This agronomic evaluation system had been effectively implemented on 196 plots. The variance in the pairwise levels of the coverage evaluation results demonstrated the robustness of the evaluation system. SoybeanTracer could indeed help to extract soybean canopy coverage and provide valuable reference for implementing soybean genotype selection, laying a useful foundation for other crops in similar scientific tasks. The source code and trained models were available at https://github.com/KLoAS-Research/SCSNet-SoybeanTracer.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:001224350500001  
Indexed Date:  
2024-05-22  
Institution Address:  
Anhui Agr Univ, Sch Informat & Artificial Intelligence, Hefei 230036, Peoples R China  
Institution Address:  
Minist Agr & Rural Affairs, Key Lab Agr Sensors, Hefei 230036, Peoples R China  
Institution Address:  
Anhui Prov Key Lab Smart Agr Technol & Equipment, Hefei 230036, Peoples R China  
Institution Address:  
Anhui Agr Univ, Sch Agron, Hefei 230036, Peoples R China  
Institution:  
Institution:  
Institution:  
Anhui Prov Key Lab Smart Agr Technol & Equipment  
Institution:

### 927. Remote Sensing Ship Target Detection Based on Feature and Region

摘要: The use of high-resolution remote sensing imagery for ship detection has a wide range of applications in military and civilian fields,such as maritime surveillance,search and rescue,and transportation.However,in highresolution optical remote sensing images,complex environment as well as arbitrary directions and variable scales of ship targets lead to poor detection accuracy.To address these limitations,a rotation detection algorithm,known as the RetinaNet-MPD,is proposed based on feature and region localization enhancement.First,the RetinaNet-MPD adds a multi-scale feature fusion module,which entirely integrates feature information at different scales and levels,to enhance the feature representation ability of feature maps at different scales.Second,a Polarized Dual-Attention Network(PDANet) module is proposed for ship target detection in a complex environment.By adding a polarization function after the attention network,the key features of the target are entirely extracted,and irrelevant information is suppressed to effectively distinguish the target from its surrounding.In addition,a Dynamic Anchor Learning(DAL) method is adopted when training the positive and negative samples to dynamically select high-quality anchors with good localization potential in the target region and improve the accuracy and precision of ship target detection.The experimental results show that RetinaNet-MPD algorithm achieved detection accuracy of 89.3% and 85.8% on the DOTA-Ship and HRSC2016 data sets,respectively.Consequently,the average detection accuracy was improved effectively compared with other existing rotating-target detection models.  
摘要:  
高分辨率遥感图像在海上监视、海上搜救、海上运输等军用和民用领域的舰船检测方面有着广泛的应用。然而高分辨率光学遥感图像舰船目标检测通常存在背景复杂、目标方向任意、尺度多变等问题,导致检测精度不高。提出一种基于特征和区域定位增强的旋转检测算法RetinaNet-MPD。通过添加一个多尺度特征融合模块,充分融合不同尺度、不同层级的特征信息,以增强不同尺度特征图的特征表示能力。针对复杂背景下的舰船目标检测,提出极化双重注意力网络,通过在注意力网络后加入极化函数,充分提取目标的关键特征,同时抑制不相关信息,以有效区分目标和背景。此外,为更准确地定位舰船目标,在对正负样本进行训练时采用一种动态锚学习方法,从而动态选择目标区域内具有良好定位潜力的高质量锚,提高舰船目标检测精度。实验结果表明,RetinaNet-MPD算法在DOTA舰船和HRSC2016数据集上的检测精度分别为89.3%和85.8%,相比现有旋转目标检测算法的检测精度有所提升。  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1000-3428  
入藏号:  
CSCD:7535123  
Indexed Date:  
2023-11-24  
Institution Address:  
湖北工业大学计算机学院, 武汉, 湖北 430000, 中国  
Institution Address:  
火箭军指挥学院勤务保障系, 武汉, 湖北 430012, 中国  
Institution Address:  
School of Computer Science,Hubei University of Technology, Wuhan, Hubei 430000, China  
Institution Address:  
Service Support Department,Rocket Army Command Academy, Wuhan, Hubei 430012, China  
Institution:  
湖北工业大学计算机学院  
Institution:  
火箭军指挥学院勤务保障系  
Institution:  
School of Computer Science,Hubei University of Technology  
Institution:  
Service Support Department,Rocket Army Command Academy

### 928. OPT-SAR-MS<SUP>2</SUP>Net: A Multi-Source Multi-Scale Siamese Network

摘要: The utilization of optical and synthetic aperture radar (SAR) multi-source data to obtain better land classification results has received increasing research attention. However, there is a large property and distributional difference between optical and SAR data, resulting in an enormous challenge to fuse the inherent correlation information to better characterize land features. Additionally, scale differences in various features in remote sensing images also influence the classification results. To this end, an optical and SAR Siamese semantic segmentation network, OPT-SAR-MS(2)Net, is proposed. This network can intelligently learn effective multi-source features and realize end-to-end interpretation of multi-source data. Firstly, the Siamese network is used to extract features from optical and SAR images in different channels. In order to fuse the complementary information, the multi-source feature fusion module fuses the cross-modal heterogeneous remote sensing information from both high and low levels. To adapt to the multi-scale features of the land object, the multi-scale feature-sensing module generates multiple information perception fields. This enhances the network's capability to learn contextual information. The experimental results obtained using WHU-OPT-SAR demonstrate that our method outperforms the state of the art, with an mIoU of 45.2% and an OA of 84.3%. These values are 2.3% and 2.6% better than those achieved by the most recent method, MCANet, respectively.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2072-4292  
入藏号:  
WOS:001247980700001  
Indexed Date:  
2024-06-22  
Institution Address:  
Acad Mil Sci AMS, Def Innovat Inst, Beijing 100000, Peoples R China  
Institution Address:  
Harbin Inst Technol, Sch Aeronaut, Harbin 150001, Peoples R China  
Institution Address:  
Shandong Inst Space Elect Technol, Yantai 250100, Peoples R China  
Institution Address:  
DFH Satellite Co Ltd, Beijing 100094, Peoples R China  
Institution:  
Acad Mil Sci AMS  
Institution:  
Institution:  
Shandong Inst Space Elect Technol  
Institution:  
DFH Satellite Co Ltd

### 929. Face hallucination using multisource references and cross-scale dual

摘要: There is an increasing interest in enhancing the quality of low-resolution (LR) facial images for various social life applications. Existing methods often use domain-specific prior knowledge, which is effective in improving the face super-resolution model's performance. However, it is challenging to obtain rich and accurate prior information from LR inputs in real-world scenarios, which can limit the robustness and generalization ability of the developed face super-resolution model. In this paper, a multisource reference-based face super-resolution Network, namely MSRNet, is proposed. Without considering the prior knowledge of faces, the network can reconstruct a LR face image with a magnitude factor of 8 under the guidance of multiple reference face images of different identities. By constructing an "appearance-alike" reference data set Face\_Ref, the designed MSRNet aims to fully exploit the local and spatially similar high frequency information between the distinct references and the current face. More specifically, to effectively combine the information from multiple references, a cross-scale and cross-space feature fusion mechanism is introduced for external and internal references, and then the enhanced local semantics are finally incorporated into the high-resolution face reconstruction. The robustness of face image super-resolution is increased compared to current correlation approaches, since it not only eliminates the need for face prior knowledge but also avoids performing alignment operations on reference faces with multiple expressions and different poses. Experimental results show that the proposed model is able to produce results for face super-resolution that are satisfying and dependable and outperforms the state-of-the-art methods in terms of visual perceptual quality and quantity evaluation.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
8  
ISSN:  
0884-8173  
eISSN:  
1098-111X  
入藏号:  
WOS:000844221800001  
Indexed Date:  
2022-09-01  
Institution Address:  
Shandong Univ Finance & Econ, Sch Comp Sci & Technol, Jinan, Peoples R China  
Institution Address:  
Linyi Univ, Sch Informat Sci & Technol, Linyi, Shandong, Peoples R China  
Institution Address:  
Univ Portsmouth, Sch Creat Technol, Portsmouth, Hants, England  
Institution Address:  
Univ Jinan, Shandong Prov Key Lab Network Based Intelligent C, Jinan, Peoples R China  
Institution Address:  
QuanCheng Lab, Jinan, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
QuanCheng Lab

### 930. MSFF-Net: Multi-scale feature fusing networks with dilated mixed

摘要: Sound Event Detection (SED) needs to identify the sound events in a recording and detect the onset and offset times of them. The former desires features with long short-term dependencies to detect sound events with different durations and the latter needs fine-grained dependency. Although our previous proposed Multi-Scale Fully Convolutional Networks (MS-FCN) uses cascaded dilated convolution to model temporal context information and multi-scale information is considered, there are two shortages to deal with: the ignorance of neighboring information and fine-grained dependencies, and neglecting intermediate-length temporal dependencies. The first shortage is caused by the skipping elements sampling mechanism of dilated convolution, by which the neighboring information and fine-grained dependencies are ignored. To overcome this shortage, the paper proposes the dilated mixed convolution module, which mixes dilated convolution and standard convolutions to capture both the fine-grained and long-term dependencies and give weight to neighboring information. The second shortage is caused by the too fast increase of temporal dependent length in cascaded dilated convolution module, which causes too much intermediate temporal information to be ignored. For this shortage, this paper proposes Dilated Temporal Pyramid Pooling module (DTPP), in which parallel dilated convolutions with multiple dilation factors are used to capture the intermediate temporal information with a proper temporal dependent length. As cascaded module has been demonstrated to be valid and efficient to model the temporal context in MS-FCN and DTPP module can capture the ignored temporal information of cascaded module, taking the advantages of both, this paper proposes the cascaded parallel module to capture richer temporal dependencies. Based on that, Multi-Scale Feature Fusion Networks (MSFF-Net) is proposed, which obtains competitive performance on three open datasets. (C)& nbsp;2021 Published by Elsevier Inc.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
8  
ISSN:  
1051-2004  
eISSN:  
1095-4333  
入藏号:  
WOS:000783022400006  
Indexed Date:  
2022-05-08  
Institution Address:  
Xidian Univ, Sch Artificial Intelligence, Xian 710071, Shaanxi, Peoples R China  
Institution Address:  
Guangdong Artificial Intelligence & Digital Econ, Guangzhou 510335, Guangdong, Peoples R China  
Institution:  
Institution:  
Guangdong Artificial Intelligence & Digital Econ

### 931. Attention-Guided Multi-Scale Feature Fusion Network for Low-Light Image

摘要: Low-light image enhancement has been an important research branch in the field of computer vision. Low-light images are characterized by poor visibility, high noise and low contrast. To improve low-light images generated in low-light environments and night conditions, we propose an Attention-Guided Multi-scale feature fusion network (MSFFNet) for low-light image enhancement for enhancing the contrast and brightness of low-light images. First, to avoid the high cost computation arising from the stacking of multiple sub-networks, our network uses a single encoder and decoder for multi-scale input and output images. Multi-scale input images can make up for the lack of pixel information and loss of feature map information caused by a single input image. The multi-scale output image can effectively monitor the error loss in the image reconstruction process. Second, the Convolutional Block Attention Module (CBAM) is introduced in the encoder part to effectively suppress the noise and color difference generated during feature extraction and further guide the network to refine the color features. Feature calibration module (FCM) is introduced in the decoder section to enhance the mapping expression between channels. Attention fusion module (AFM) is also added to capture contextual information, which is more conducive to recovering image detail information. Last, the cascade fusion module (CFM) is introduced to effectively combine the feature map information under different perceptual fields. Sufficient qualitative and quantitative experiments have been conducted on a variety of publicly available datasets, and the proposed MSFFNet outperforms other low-light enhancement methods in terms of visual effects and metric scores.  
Web of Science 核心合集中的 "被引频次":  
5  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
1662-5218  
入藏号:  
WOS:000772100800001  
PubMed ID:  
35308314  
Indexed Date:  
2022-03-30  
Institution Address:  
Shandong Technol Business Univ, Coll Elect & Commun Engn, Yantai, Peoples R China  
Institution Address:  
Inst Comp Technol ICT, Inst Network Technol, Yantai, Peoples R China  
Institution Address:  
Shandong Univ Finance & Econ, Sch Comp Sci & Technol, Jinan, Peoples R China  
Institution:  
Institution:  
Inst Comp Technol ICT  
Institution:

### 932. Direction-aware multi-branch attention and Gaussian label assignment for

摘要: A large number of objects of different scales appear aggregative distribution in remote sensing images, which brings two challenges to remote sensing object detection. The first challenge arises from the presence of objects with large scale differences within the same scene, which complicates the detector's ability to balance the detection performance between large and small objects. The second challenge is that the distribution of dense objects on remote sensing images with complex backgrounds can reduce the feature representation capability of the network, leading to false detections and missed detections of objects. To address the two challenges present in remote sensing object detection, we propose a new object detection network DA-GLANet (Direction-Aware Multi-Branch Attention and Gaussian Label Assignment Balance Net). For the first challenge, we designed a multi-scale feature fusion module to fully utilize semantic information at different scales, thereby improving the detection accuracy of multi-scale objects. Subsequently, to tackle the second challenge, we proposed a multi-branch attention module with direction-aware capabilities. This module can enhance the network's feature representation of objects of interest by generating attention maps that are both direction-aware and position-sensitive. In addition, we propose a label assignment strategy based on Gaussian distribution to address the imbalance of positive and negative labels caused by the complexity of remote sensing images. These methods work synergistically to improve aggregative object detection accuracy. We have conducted extensive experiments on the DOTA dataset, Levir dataset, and WHU buildings change detection dataset to demonstrate the superiority of our method. On the DOTA dataset, our method improves mAP by 0.4%, 0.3%, and 0.2% compared to the top three methods. On the Levir dataset, our method improves mAP by 2.5%, 2.2%, and 1.6% compared to the top three methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:001283150800001  
Indexed Date:  
2024-08-10  
Institution Address:  
Southwest Jiaotong Univ, Fac Geosci & Environm Engn, 999 Xian Rd, Chengdu 611756, Sichuan, Peoples R China  
Institution:

### 933. An efficient re-parameterization feature pyramid network on YOLOv8 to

摘要: In the field of steel production, the detection of steel surface defects is one of the most important guarantees for the quality of steel production. In the process of defect detection, there are problems regarding the noise of the acquisition background, the scale of defects, and the detection speed. At present, in the face of complex steel surface defects, realizing efficient real-time steel surface defect detection has become a difficult problem. In this paper, we propose a lightweight and efficient real-time defect detection method, LDE-YOLO, based on YOLOv8. First, we propose a lightweight multi-scale feature extraction module, LighterMSMC, which not only achieves a lightweight backbone network, but also effectively guarantees the long range dependence of the features, so as to realize multi-scale feature extraction more efficiently. Secondly, we propose lightweight re-parameterized feature pyramid, DE-FPN, in which the sparse patterns of the overall features and the detailed features of the local features are efficiently captured by the DE-Block, and then efficiently fused by the PAN feature fusion structure. Finally, we propose Efficient Head, which lightens the model by group convolution while its improves the diagonal correlation of the feature maps on some specific datasets, thus enhancing the detection performance. Our proposed LDE-YOLO obtains 80.8 mAP and 75.5 FPS on NEU-DET , 80.5 mAP and 75.5 FPS on GC10-DET. It obtains 2.5 mAP and 4.7 mAP enhancement compared to the baseline model, and the detection speed is also improved by 10.4 FPS, while in terms of the number of floating point operations and parameters of the model reduced by 60.2% and 49.1%, which is sufficient to illustrate its lightweight effectiveness and realize an efficient real-time steel surface defect detection model.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:001353681600001  
Indexed Date:  
2024-11-19  
Institution Address:  
Zhejiang Univ Sci & Technol, Sch Informat & Elect Engn, 318 Liuhe Rd, Hangzhou 310023, Zhejiang, Peoples R China  
Institution:

### 934. LungSeek: 3D Selective Kernel residual network for pulmonary nodule

摘要: Early detection and diagnosis of pulmonary nodules is the most promising way to improve the survival chances of lung cancer patients. This paper proposes an automatic pulmonary cancer diagnosis system, LungSeek. LungSeek is mainly divided into two modules: (1) Nodule detection, which detects all suspicious nodules from computed tomography (CT) scan; (2) Nodule Classification, classifies nodules as benign or malignant. Specifically, a 3D Selective Kernel residual network (SK-ResNet) based on the Selective Kernel Network and 3D residual network is located. A deep 3D region proposal network with SK-ResNet is designed for detection of pulmonary nodules while a multi-scale feature fusion network is designed for the nodule classification. Both networks use the SK-Net module to obtain different receptive field information, thereby effectively learning nodule features and improving diagnostic performance. Our method has been verified on the luna16 data set, reaching 89.06, 94.53% and 97.72% when the average number of false positives is 1, 2 and 4, respectively. Meanwhile, its performance is better than the state-of-the-art method and other similar networks and experienced doctors. This method has the ability to adaptively adjust the receptive field according to multiple scales of the input information, so as to better detect nodules of various sizes. The framework of LungSeek based on 3D SK-ResNet is proposed for nodule detection and nodule classification from chest CT. Our experimental results demonstrate the effectiveness of the proposed method in the diagnosis of pulmonary nodules.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
15  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
15  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:000749067200001  
PubMed ID:  
35103029  
Indexed Date:  
2022-02-06  
Institution Address:  
Wuhan Univ Sci & Technol, Coll Comp Sci & Technol, Wuhan 430081, Peoples R China  
Institution Address:  
Hubei Prov Key Lab Intelligent Informat Proc & Re, Wuhan, Peoples R China  
Institution:  
Institution:  
Hubei Prov Key Lab Intelligent Informat Proc & Re

### 935. ASCEND-UNet: An Improved UNet Configuration Optimized for Rural

摘要: Different types of rural settlement agglomerations have been formed and mixed in space during the rural revitalization strategy implementation in China. Discriminating them from remote sensing images is of great significance for rural land planning and living environment improvement. Currently, there is a lack of automatic methods for obtaining information on rural settlement differentiation. In this paper, an improved encoder-decoder network structure, ASCEND-UNet, was designed based on the original UNet. It was implemented to segment and classify dispersed and clustered rural settlement buildings from high-resolution satellite images. The ASCEND-UNet model incorporated three components: firstly, the atrous spatial pyramid pooling (ASPP) multi-scale feature fusion module was added into the encoder, then the spatial and channel squeeze and excitation (scSE) block was embedded at the skip connection; thirdly, the hybrid dilated convolution (HDC) block was utilized in the decoder. In our proposed framework, the ASPP and HDC were used as multiple dilated convolution blocks to expand the receptive field by introducing a series of dilated rate convolutions. The scSE is an attention mechanism block focusing on features both in the spatial and channel dimension. A series of model comparisons and accuracy assessments with the original UNet, PSPNet, DeepLabV3+, and SegNet verified the effectiveness of our proposed model. Compared with the original UNet model, ASCEND-UNet achieved improvements of 4.67%, 2.80%, 3.73%, and 6.28% in precision, recall, F1-score and MIoU, respectively. The contributions of HDC, ASPP, and scSE modules were discussed in ablation experiments. Our proposed model obtained more accurate and stable results by integrating multiple dilated convolution blocks with an attention mechanism. This novel model enriches the automatic methods for semantic segmentation of different rural settlements from remote sensing images.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001311680300001  
PubMed ID:  
39275364  
Indexed Date:  
2024-09-17  
Institution Address:  
Zhejiang A&F Univ, Coll Math & Comp Sci, Hangzhou 311300, Peoples R China  
Institution Address:  
State Forestry & Grassland Adm Forestry Sensing Te, Key Lab, Hangzhou 311300, Peoples R China  
Institution Address:  
Key Lab Forestry Intelligent Monitoring & Informat, Hangzhou 311300, Peoples R China  
Institution:  
Institution:  
State Forestry & Grassland Adm Forestry Sensing Te  
Institution:  
Key Lab Forestry Intelligent Monitoring & Informat

### 936. Mf-net: multi-feature fusion network based on two-stream extraction and

摘要: Due to the increasing sophistication of face forgery techniques, the images generated are becoming more and more realistic and difficult for human eyes to distinguish. These face forgery techniques can cause problems such as fraud and social engineering attacks in facial recognition and identity verification areas. Therefore, researchers have worked on face forgery detection studies and have made significant progress. Current face forgery detection algorithms achieve high detection accuracy within-dataset. However, it is difficult to achieve satisfactory generalization performance in cross-dataset scenarios. In order to improve the cross-dataset detection performance of the model, this paper proposes a multi-feature fusion network based on two-stream extraction and multi-scale enhancement. First, we design a two-stream feature extraction module to obtain richer feature information. Secondly, the multi-scale feature enhancement module is proposed to focus the model more on information related to the current sub-region from different scales. Finally, the forgery detection module calculates the overlap between the features of the input image and real images during the training phase to determine the forgery regions. The method encourages the model to mine forgery features and learns generic and robust features not limited to a particular feature. Thus, the model achieves high detection accuracy and performance. We achieve the AUC of 99.70% and 90.71% on FaceForensics++ and WildDeepfake datasets. The generalization experiments on Celeb-DF-v2 and WildDeepfake datasets achieve the AUC of 80.16% and 65.15%. Comparison experiments with multiple methods on other benchmark datasets confirm the superior generalization performance of our proposed method while ensuring model detection accuracy. Our code can be found at: https://github.com/1241128239/MFNet.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2199-4536  
eISSN:  
2198-6053  
入藏号:  
WOS:001351565500004  
Indexed Date:  
2024-11-21  
Institution Address:  
Yunnan Univ, Engn Res Ctr Cyberspace, Kunming 650091, Peoples R China  
Institution Address:  
Wroclaw Univ Sci & Technol, Informat & Commun Technol, Wroclaw, Poland  
Institution:  
Institution:

### 937. SC-Lite: An Efficient Lightweight Model for Real-Time X-Ray Security

摘要: Efficient and rapid analysis of X-ray images is crucial for detecting contraband, including weapons and explosives, at transportation hubs. While existing models like YOLOv8 offer high performance, they often fall short in real-time scenarios due to slow processing speeds. This study proposes SC-Lite, an improved model based on YOLOv8, designed to improve real-time detection of prohibited items in environments with limited computational power, such as embedded systems. We propose a novel acceleration module, the CSPNet Faster Convolution Network Module (C2F\_FM), which reduces computational redundancy and optimizes memory usage by focusing standard convolutions on key channels, thereby enhancing efficiency. Furthermore, the Adaptation-BiFPN module improves multi-scale feature fusion with dynamic, content-aware weighting, increasing both the sensitivity and accuracy of detections. Additionally, the LAMP pruning strategy is utilized to optimize the model for deployment in resource-limited environments, significantly reducing the complexity of its architecture. After extensive experimental validation, SC-Lite outperforms all mainstream models in terms of real-time performance and detection accuracy, and the model outperforms existing methods in all metrics. The model achieved reductions of 72%, 62%, and 70% in parameters, computation, and model size, respectively. On the LSIray dataset, SC-Lite boosted the mean average precision (mAP) by 0.3% to 94.3% and enhanced the frame rate from 87 FPS to 136 FPS, an improvement of 49 FPS. For the OPIXray dataset, SC-Lite increased the mAP by 1.5% to 94.8%, with the frame rate rising from 344 FPS to 400 FPS, an improvement of 56 FPS. This study provides an innovative solution to improve the real-time performance and accuracy of security inspection systems.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001283754200001  
Indexed Date:  
2024-08-10  
Institution Address:  
Zhengzhou Univ Light Ind, Sch Comp Sci & Technol, Zhengzhou 450002, Henan, Peoples R China  
Institution:

### 938. Collaborative Multi-Metadata Fusion to Improve the Classification of

摘要: Computed tomography (CT) images are the most commonly used radiographic imaging modality for detecting and diagnosing lumbar diseases. Despite many outstanding advances, computer-aided diagnosis (CAD) of lumbar disc disease remains challenging due to the complexity of pathological abnormalities and poor discrimination between different lesions. Therefore, we propose a Collaborative Multi-Metadata Fusion classification network (CMMF-Net) to address these challenges. The network consists of a feature selection model and a classification model. We propose a novel Multi-scale Feature Fusion (MFF) module that can improve the edge learning ability of the network region of interest (ROI) by fusing features of different scales and dimensions. We also propose a new loss function to improve the convergence of the network to the internal and external edges of the intervertebral disc. Subsequently, we use the ROI bounding box from the feature selection model to crop the original image and calculate the distance features matrix. We then concatenate the cropped CT images, multiscale fusion features, and distance feature matrices and input them into the classification network. Next, the model outputs the classification results and the class activation map (CAM). Finally, the CAM of the original image size is returned to the feature selection network during the upsampling process to achieve collaborative model training. Extensive experiments demonstrate the effectiveness of our method. The model achieved 91.32% accuracy in the lumbar spine disease classification task. In the labelled lumbar disc segmentation task, the Dice coefficient reaches 94.39%. The classification accuracy in the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) reaches 91.82%.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0278-0062  
eISSN:  
1558-254X  
入藏号:  
WOS:001122030500021  
PubMed ID:  
37432809  
Indexed Date:  
2023-12-28  
Institution Address:  
Shandong Univ, Sch Software, Jinan 250100, Peoples R China  
Institution:

### 939. An Adaptive Region Proposal Network With Progressive Attention

摘要: Two-stage detectors, which consist of the multi-scale feature representations and the prediction of region proposal boxes, have been recognized as an effective paradigm for tiny object detection in Unmanned Aerial Vehicle (UAV) images. Although most previous methods primarily concentrated on developing efficient feature fusion strategies within the feature pyramid network (FPN), few studies elaborated on improving the performance of region proposal network (RPN). Conventional RPNs exhibit two key weaknesses in the majority of existing two-stage object detection approaches. Firstly, the quality of proposal boxes generated by the RPN is heavily reliant on rich feature representations extracted from the FPN backbone. Secondly, the fixed number of generated proposal boxes limits adaptability to the distribution of tiny person objects. To mitigate the aforementioned problems, in this paper we propose a novel adaptive region proposal network (ARPN) to improve the quality of the proposal boxes and generate particularly compact yet accurate proposal boxes. On one hand, a progressive attention mechanism is devised to make the ARPN focus more on prospective object regions, where a series of multi-scale front attention modules (FAM) are applied to coarsely filter out most of irrelevant background areas and a group of top-to-bottom back attention modules (BAM) aid the ARPN to finely pinpoint tiny objects of interest in a coarse-to-fine manner. On the other hand, a mini-density map, which is inspired by the philosophy of crowd counting, is elaborately designed to adaptively determine the number of region proposal boxes. This approach significantly reduces redundancy while maintaining high-quality proposal boxes. Extensive experiments verify the superiority of proposed ARPN and show obvious improvement over other competitors in terms of two performance indicators of average precision (AP) and average recall (AR). The code will be available at https://github.com/kbzhang0505/ARPN.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:001241605300053  
Indexed Date:  
2024-08-30  
Institution Address:  
Xian Polytech Univ, Shaanxi Key Lab Clothing Intelligence, Sch Comp Sci, Xian 710048, Peoples R China  
Institution Address:  
Xian Polytech Univ, Sch Elect & Informat, Xian 710048, Peoples R China  
Institution Address:  
Xidian Univ, State Key Lab Integrated Serv Networks, Xian 710071, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Chongqing Key Lab Image Cognit, Chongqing 400065, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 940. Lightweight Target Detection for Coal and Gangue Based on Improved

摘要: The detection of coal and gangue is an essential part of intelligent sorting. A lightweight coal and gangue detection algorithm based on You Only Look Once version 5s (Yolov5s) is proposed for the current coal and gangue target detection algorithm with the low accuracy of small target detection, high model complexity, and sizeable computational memory consumption. Firstly, we build a new convolutional block based on the Funnel Rectified Linear Unit (FReLU) activation function and apply it to the original Yolov5s network so that the model adaptively captures local contextual information of the image. Secondly, the neck of the original network is redesigned to improve the detection accuracy of small samples by adding a small target detection head to achieve multi-scale feature fusion. Next, some of the standard convolution modules in the original network are replaced with Depthwise Convolution (DWC) and Ghost Shuffle Convolution (GSC) modules to build a lightweight feature extraction network while ensuring the model detection accuracy. Finally, an efficient channel attention (ECA) module is embedded in the backbone of the lightweight network to facilitate accurate localization of the prediction region by improving the information interaction of the model with the channel features. In addition, the importance of each component is fully demonstrated by ablation experiments and visualization analysis comparison experiments. The experimental results show that the mean average precision (mAP) and the model size of our proposed model reach 0.985 and 4.9 M, respectively. The mAP is improved by 0.6%, and the number of parameters is reduced by 72.76% compared with the original Yolov5s network. The improved algorithm has higher localization and recognition accuracy while significantly reducing the number of floating-point calculations and of parameters, reducing the dependence on hardware, and providing a specific reference basis for deploying automated underground gangue sorting.  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
5  
eISSN:  
2227-9717  
入藏号:  
WOS:000977704700001  
Indexed Date:  
2023-05-14  
Institution Address:  
Anhui Univ Sci & Technol, Sch Elect & Informat Engn, Huainan 232001, Peoples R China  
Institution:

### 941. A Generative Adversarial Network with Spatial Attention Mechanism for

摘要: The acquisition of building structures has broad applications across various fields. However, existing methods for inferring building structures predominantly depend on manual expertise, lacking sufficient automation. To tackle this challenge, we propose a building structure inference network that utilizes UAV remote sensing images, with the PIX2PIX network serving as the foundational framework. We enhance the generator by incorporating an additive attention module that performs multi-scale feature fusion, enabling the combination of features from diverse spatial resolutions of the feature map. This modification enhances the model's capability to emphasize global relationships during the mapping process. To ensure the completeness of line elements in the generator's output, we design a novel loss function based on the Hough transform. A line penalty term is introduced that transforms the output of the generator and ground truth to the Hough domain due to the original loss function's inability to effectively constrain the completeness of straight-line elements in the generated results in the spatial domain. A dataset of the appearance features obtained from UAV remote sensing images and the internal floor plan structure is made. Using UAV remote sensing images of multi-story residential buildings, high-rise residential buildings, and office buildings as test collections, the experimental results show that our method has better performance in inferring a room's layout and the locations of load-bearing columns, achieving an average improvement of 11.2% and 21.1% over PIX2PIX in terms of the IoU and RMSE, respectively.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001077750600001  
Indexed Date:  
2023-10-18  
Institution Address:  
Harbin Inst Technol, Sch Elect & Informat Engn, Harbin 150006, Peoples R China  
Institution Address:  
Acad Mil Sci, Inst Def Engn, Beijing 100036, Peoples R China  
Institution:  
Institution:  
Acad Mil Sci

### 942. Contrastive Self-Supervised Two-Domain Residual Attention Network with

摘要: Hyperspectral images can assist change-detection methods in precisely identifying differences in land cover in the same region at different observation times. However, the difficulty of labeling hyperspectral images restricts the number of training samples for supervised change-detection methods, and there are also complex real influences on hyperspectral images, such as noise and observation directions. Furthermore, current deep-learning-based change-detection methods ignore the feature reusage from receptive fields with different scales and cannot effectively suppress unrelated spatial-spectral dependencies globally. To better handle these issues, a contrastive self-supervised two-domain residual attention network (TRAMNet) with a random augmentation pool is proposed for hyperspectral change detection. The contributions of this article are summarized as follows. (1) To improve the feature extraction from hyperspectral images with random Gaussian noise and directional information, a contrastive learning framework with a random data augmentation pool and a soft contrastive loss function (SCLF) is proposed. (2) The multi-scale feature fusion module (MFF) is provided to achieve feature reusage from different receptive fields. (3) A two-domain residual attention (TRA) block is designed to suppress irrelated change information and extract long-range dependencies from both spectral and spatial domains globally. Extensive experiments were carried out on three real datasets. The results show that the proposed TRAMNet can better initialize the model weights for hyperspectral change-detection task and effectively decrease the need for training samples. The proposed method outperforms most existing hyperspectral change-detection methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
eISSN:  
2072-4292  
入藏号:  
WOS:001045573200001  
Indexed Date:  
2023-08-23  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, 20 Datun Rd, Beijing 100101, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, 3 Datun Rd, Beijing 100101, Peoples R China  
Institution:  
Institution:

### 943. A unified architecture for super-resolution and segmentation of remote

摘要: The resolution of the image has an important impact on the accuracy of segmentation. Integrating super- resolution (SR) techniques in the semantic segmentation of remote sensing images contributes to the improvement of precision and accuracy, especially when the images are blurred. In this paper, a novel and efficient SR semantic segmentation network (SRSEN) is designed by taking advantage of the similarity between SR and segmentation tasks in feature processing. SRSEN consists of the multi-scale feature encoder, the SR fusion decoder, and the multi-path feature refinement block, which adaptively establishes the feature associations between segmentation and SR tasks to improve the segmentation accuracy of blurred images. Experiments show that the proposed method achieves higher segmentation accuracy on fuzzy images compared to state-of-the-art models. Specifically, the mIoU of the proposed SRSEN is 3%-6% higher than other state-of-the-art models on low-resolution LoveDa, Vaihingen, and Potsdam datasets.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0141-9382  
eISSN:  
1872-7387  
入藏号:  
WOS:001292231600001  
Indexed Date:  
2024-08-21  
Institution Address:  
Linyi Univ, Dept Informat Sci & Engn, Linyi 276000, Peoples R China  
Institution Address:  
Shandong Univ, Dept Control Sci & Engn, Jinan 250061, Peoples R China  
Institution Address:  
Shandong Res Inst Ind Technol, Jinan 250013, Peoples R China  
Institution Address:  
YanCheng Inst Technol, Dept Informat Sci & Engn, Yancheng 224000, Peoples R China  
Institution:  
Institution:  
Institution:  
Shandong Res Inst Ind Technol  
Institution:

### 944. Multi-Scale Fusion Uncrewed Aerial Vehicle Detection Based on RT-DETR

摘要: With the rapid development of science and technology, uncrewed aerial vehicle (UAV) technology has shown a wide range of application prospects in various fields. The accuracy and real-time performance of UAV target detection play a vital role in ensuring safety and improving the work efficiency of UAVs. Aimed at the challenges faced by the current UAV detection field, this paper proposes the Gathering Cascaded Dilated DETR (GCD-DETR) model, which aims to improve the accuracy and efficiency of UAV target detection. The main innovations of this paper are as follows: (1) The Dilated Re-param Block is creatively applied to the dilatation-wise Residual module, which uses the large kernel convolution and the parallel small kernel convolution together and fuses the feature maps generated by multi-scale perception, greatly improving the feature extraction ability, thereby improving the accuracy of UAV detection. (2) The Gather-and-Distribute mechanism is introduced to effectively enhance the ability of multi-scale feature fusion so that the model can make full use of the feature information extracted from the backbone network and further improve the detection performance. (3) The Cascaded Group Attention mechanism is innovatively introduced, which not only saves the computational cost but also improves the diversity of attention by dividing the attention head in different ways, thus enhancing the ability of the model to process complex scenes. In order to verify the effectiveness of the proposed model, this paper conducts experiments on multiple UAV datasets of complex scenes. The experimental results show that the accuracy of the improved RT-DETR model proposed in this paper on the two UAV datasets reaches 0.956 and 0.978, respectively, which is 2% and 1.1% higher than that of the original RT-DETR model. At the same time, the FPS of the model is also improved by 10 frames per second, which achieves an effective balance between accuracy and speed.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
7  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
2079-9292  
入藏号:  
WOS:001209920700001  
Indexed Date:  
2024-05-04  
Institution Address:  
Beijing Informat Sci & Technol Univ, Comp Sch, Beijing 100101, Peoples R China  
Institution:

### 945. Point cloud completion using multiscale feature fusion and

摘要: Raw point clouds obtained from real-world scanning are always incomplete and ununiformly distributed, which would result in structural losses of object shapes and bring about difficulties in further high-level 3D vision tasks. Therefore, a learning-based method called CRA-Net is proposed in this paper to repair partial point clouds and predict complete object shapes. Compared with most existing networks that only leverage global features, CRA-Net successfully utilizes local features to restore clearer details of object shapes with low instability. First, we propose an adaptive neighborhood query method that is able to adjust query centers and radiuses to cover different object shapes and acquire balanced local regions. Second, we build a parallel encoder to extract multiscale features from the input. Third, we design a cross-regional attention module based on graph attention network. It quantifies underlying relationships among all the local features under certain conditions interpreted by global features. Based on such relationships, each conditional local feature vector is able to search across the regions and selectively absorb other local features. Fourth, we design a coarse decoder to collect these cross region features and generate the skeleton of complete point cloud. Finally, we refine the coarse point cloud by comparing it with the input, and up sample it using folding-based layers.  
Our network is first trained and tested on manually made partial-complete point clouds pairs generated by the scanning process of a virtual LiDAR on eight categories of objects. Then it is tested on real-world point clouds of indoor and outdoor scenes. Compared with existing representative methods, our CRA-Net always restores the most accurate point clouds with the clearest details.  
(c) 2021 Elsevier B.V. All rights reserved.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
16  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
18  
ISSN:  
0262-8856  
eISSN:  
1872-8138  
入藏号:  
WOS:000658385700004  
Indexed Date:  
2021-06-25  
Institution Address:  
Shanghai Jiao Tong Univ, Shanghai 200240, Peoples R China  
Institution:

### 946. Low-Light Image Enhancement Network Based on Recursive Network

摘要: In low-light environments, image acquisition devices do not obtain sufficient light sources, resulting in low brightness and contrast of images, which poses a great obstacle for other computer vision tasks to be performed. To enable other vision tasks to be performed smoothly, it is essential to enhance the research on low-light image enhancement algorithms. In this article, a multi-scale feature fusion image enhancement network based on recursive structure is proposed. The network uses a dual attention module-Convolutional Block Attention Module. It was abbreviated as CBAM, which includes two attention mechanisms: channel attention and spatial attention. To extract and fuse multi-scale features, we extend the U-Net model using the inception model to form the Multi-scale inception U-Net Module or MIU module for short. The learning of the whole network is divided into T recursive stages, and the input of each stage is the original low-light image and the inter-mediate estimation result of the output of the previous recursion. In the t-th recursion, CBAM is first used to extract channel feature information and spatial feature information to make the network focus more on the low-light region of the image. Next, the MIU module fuses features from three different scales to obtain inter-mediate enhanced image results. Finally, the inter-mediate enhanced image is stitched with the original input image and fed into the t + 1th recursive iteration. The inter-mediate enhancement result provides higher-order feature information, and the original input image provides lower-order feature information. The entire network outputs the enhanced image after several recursive cycles. We conduct experiments on several public datasets and analyze the experimental results subjectively and objectively. The experimental results show that although the structure of the network in this article is simple, the method in this article can recover the details and increase the brightness of the image better and reduce the image degradation compared with other methods.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
7  
ISSN:  
1662-5218  
入藏号:  
WOS:000776745100001  
PubMed ID:  
35360834  
Indexed Date:  
2022-04-22  
Institution Address:  
Shandong Technol & Business Univ, Coll Elect & Commun Engn, Yantai, Peoples R China  
Institution Address:  
Inst Network Technol, ICT, Yantai, Peoples R China  
Institution Address:  
Shandong Technol & Business Univ, Sch Comp Sci & Technol, Yantai, Peoples R China  
Institution Address:  
Shandong Technol & Business Univ, Sch Comp Sci & Technol, Yantai, Peoples R China  
Institution:  
Institution:  
Inst Network Technol  
Institution:  
Institution:

### 947. Multiscale and Hierarchical Feature-Aggregation Network for Segmenting

摘要: We propose an encoder-decoder architecture using wide and deep convolutional layers combined with different aggregation modules for the segmentation of medical images. Initially, we obtain a rich representation of features that span from low to high levels and from small to large scales by stacking multiple k x k kernels, where each k x k kernel operation is split into k x 1 and 1 x k convolutions. In addition, we introduce two feature-aggregation modules-multiscale feature aggregation (MFA) and hierarchical feature aggregation (HFA)-to better fuse information across end-to-end network layers. The MFA module progressively aggregates features and enriches feature representation, whereas the HFA module merges the features iteratively and hierarchically to learn richer combinations of the feature hierarchy. Furthermore, because residual connections are advantageous for assembling very deep networks, we employ an MFA-based long residual connections to avoid vanishing gradients along the aggregation paths. In addition, a guided block with multilevel convolution provides effective attention to the features that were copied from the encoder to the decoder to recover spatial information. Thus, the proposed method using feature-aggregation modules combined with a guided skip connection improves the segmentation accuracy, achieving a high similarity index for ground-truth segmentation maps. Experimental results indicate that the proposed model achieves a superior segmentation performance to that obtained by conventional methods for skin-lesion segmentation, with an average accuracy score of 0.97 on the ISIC-2018, PH2, and UFBA-UESC datasets.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
5  
eISSN:  
1424-8220  
入藏号:  
WOS:000794738400001  
PubMed ID:  
35591129  
Indexed Date:  
2022-05-20  
Institution Address:  
Chosun Univ, Dept Informat & Commun Engn, Gwangju 61452, South Korea  
Institution Address:  
Hankuk Univ Foreign Studies, Div Comp & Elect Syst Engn, Yongin 17035, South Korea  
Institution:  
Institution:

### 948. Patch-Based Semantically Enhanced Network for IR Dim and Small Targets

摘要: The task of background suppression in infrared small-target scenarios aims to eliminate irregular noisy backgrounds while preserving targets with high-frequency features. In infrared small-target scenes at long distances, the backgrounds become complex and the target features are degraded, highlighting a significant disparity between the detailed and realistic background and the limited features of the targets. To address these challenges, we propose a patch-based semantically enhanced generative adversarial network (GAN) named PSEnet for background suppression in infrared small-target scenarios. First, we introduce a patch-scale GAN that allows the model to concentrate on local background suppression. This shift from a global to local perspective simplifies the complexity of background suppression. Second, we employ the PSE module consisting multiscale dilated convolution and adaptive weight fusion to extract local semantic information. Third, by segmenting the infrared image into smaller patches and resampling them, we create a more balanced dataset for adversarial training. Experimental results demonstrate that the proposed algorithm significantly improves the signal-to-noise ratio of dim and small targets, reduces the missing detection rate, and achieves a precision of almost 91%. In conclusion, this approach effectively uses GANs for background suppression in complex environments.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001225641400012  
Indexed Date:  
2024-06-07  
Institution Address:  
East China Univ Sci & Technol, Key Lab Smart Mfg Energy Chem Proc, Minist Educ, Shanghai 200237, Peoples R China  
Institution Address:  
East China Univ Sci & Technol, Dept Comp Sci & Engn, Shanghai 200237, Peoples R China  
Institution Address:  
Shanghai Aerosp Control Technol Inst, Shanghai 201109, Peoples R China  
Institution Address:  
China Aerosp Sci & Technol Corp, Res & Dev Ctr Infrared Detect Technol, Shanghai 201109, Peoples R China  
Institution:  
Institution:  
Institution:  
Shanghai Aerosp Control Technol Inst  
Institution:

### 949. Salience-Aware Face Presentation Attack Detection via Deep Reinforcement

摘要: In this paper, we propose a salience-aware face presentation attack detection (SAFPAD) approach, which takes advantage of deep reinforcement learning to exploit the salient local part information in face images. Most existing deep face presentation attack detection approaches extract features from the entire image or several fixed regions. However, the discriminative information beneficial for presentation attack detection is unevenly distributed in the image due to the illumination and presentation attack instrument variation, so treating all regions equally fails to highlight the most discriminative information which is important for more accurate and robust face presentation attack detection. To address this, we propose to identify the discriminative salient parts using deep reinforcement learning and focus on them to alleviate the adverse effects of redundant information in the face images. We fuse the high-level features and the local features which guide the policy network to exploit discriminative patches and assist the classification network to predict more accurate results. We jointly train the SAFPAD model with deep reinforcement learning to generate salient locations. Extensive experiments on five public datasets demonstrate that our approach achieves very competitive performance due to the concentrated employment of salient local information.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
13  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
14  
ISSN:  
1556-6013  
eISSN:  
1556-6021  
入藏号:  
WOS:000748395300005  
Indexed Date:  
2022-02-06  
Institution Address:  
Tsinghua Shenzhen Int Grad Sch, Shenzhen 518055, Peoples R China  
Institution Address:  
Tsinghua Univ, Beijing Res Ctr Informat Sci & Technol BNRist, Dept Automat, Beijing 100084, Peoples R China  
Institution:  
Institution:

### 950. Steel Sheet Counting From an Image With a Two-Stream Network

摘要: Steel sheets play a pivotal role in a wide range of industrial processes, including the production of ships and vehicles, as well as the construction of buildings and bridges. Meanwhile, counting steel sheets accurately is essential for effective production management in factories. However, manual counting of large numbers of stacked steel sheets can lead to visual vertigo, resulting in inaccurate counts. Moreover, physical methods such as weighing are also labor-intensive and inconvenient. Fortunately, advancements in computer vision technology have opened up new possibilities for efficient steel sheet counting. Nevertheless, implementing an automatic counting method encounters challenges due to the limited texture features present in steel sheets. In this article, we present a novel approach to count steel sheets from a captured image. To the best of our knowledge, this is the pioneering work that addresses this problem using a computational approach. We make the following contributions. First, we construct a comprehensive steel sheet dataset that contains steel sheet images with corresponding manually annotated dots. Second, we propose a novel network, called TSNet, which effectively extracts features from both the RGB image and its gradient map for precise steel sheet counting. Third, we conduct extensive experiments to evaluate the effectiveness of our proposed method and demonstrate its superiority over carefully chosen baselines from state-of-the-art counting methods.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001442898800046  
Indexed Date:  
2025-03-21  
Institution Address:  
Dalian Univ Technol, Sch Comp Sci & Technol, Key Lab Social Comp & Cognit Intelligence, Minist Educ, Dalian 116024, Peoples R China  
Institution Address:  
Natl Univ Singapore, Dept Elect & Comp Engn, Singapore 117583, Singapore  
Institution Address:  
Chinese Acad Sci, Inst Automat, Beijing 100190, Peoples R China  
Institution Address:  
Dalian Univ Technol, Sch Control Sci & Engn, Dalian 116024, Peoples R China  
Institution Address:  
Dalian Univ, Sch Software Engn, Natl & Local Joint Engn Lab Comp Aided Design, Dalian 116622, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 951. Mechanical Assembly Monitoring Method Based on Semi-Supervised Semantic

摘要: Semantic segmentation of assembly images is to recognize the assembled parts and find wrong assembly operations. However, the training of supervised semantic segmentation requires a large amount of labeled data, which is time-consuming and laborious. Moreover, the sizes of mechanical assemblies are not uniform, leading to low segmentation accuracy of small-target objects. This study proposes an adversarial learning network for semi-supervised semantic segmentation of mechanical assembly images (AdvSemiSeg-MA). A fusion method of ASFF multiscale output is proposed, which combines the outputs of different dimensions of ASFF into one output. This fusion method can make full use of the high-level semantic features and low-level fine-grained features, which helps to improve the segmentation accuracy of the model for small targets. Meanwhile, the multibranch structure RFASPP module is proposed, which enlarges the receptive field and ensures the target object is close to the center of the receptive field. The CoordConv module is introduced to allow the convolution to perceive spatial position information, thus enabling the semantic segmentation network to be position-sensitive. In the discriminator network, spectral normalization is introduced. The proposed method obtains state-of-art results on the synthesized assembly depth image dataset and performs well on actual assembly RGB image datasets.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
eISSN:  
2076-3417  
入藏号:  
WOS:000914227600001  
Indexed Date:  
2023-02-06  
Institution Address:  
Qingdao Univ Technol, Sch Mech & Automot Engn, Qingdao 266000, Peoples R China  
Institution Address:  
Qingdao Univ Technol, Key Lab Ind Fluid Energy Conservat & Pollut Contro, Minist Educ, Qingdao 266000, Peoples R China  
Institution:  
Institution:

### 952. Surface Defect Detection of Power Equipment Using Adaptive Receptive

摘要: For the detection of defects such as icing,rust,and contamination of power equipment in substations,a novel adaptive receptive field network (ARFN) is proposed,in which an adaptive receptive field module (ARFM) combined with the attention mechanism can effectively fuse multi-scale features.Considering the small sample learning attribute of defect detection,a power equipment surface defect simulation data synthesis method based on real texture is also proposed.The experimental results on the simulation dataset show that the network has high detection accuracy for surface defects across devices,while having advantages such as small size and fast operation speed.  
摘要:  
针对变电站电力设备覆冰、锈蚀、污秽等缺陷检测问题,提出了一种新的自适应感受野网络,其中结合注意力机制的自适应感受野模块可对多尺度特征进行有效融合。考虑到缺陷检测的小样本学习属性,还提出了一种基于真实纹理的电力设备表面缺陷仿真数据合成方法。在仿真数据集上的实验结果表明,该网络对跨设备表面缺陷的检测精度较高,同时具有体积小、运算速度快等优点。  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1004-731X  
入藏号:  
CSCD:7508359  
Indexed Date:  
2023-10-22  
Institution Address:  
西北工业大学计算机学院, 西安, 陕西 710072, 中国  
Institution Address:  
国网浙江省电力有限公司信息通信分公司, 杭州, 浙江 310020, 中国  
Institution Address:  
School of Computer Science,Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China  
Institution Address:  
Information & Telecommunication Branch,State Grid Zhejiang Electric Power Co.,Ltd., Hangzhou, Zhejiang 310020, China  
Institution:  
西北工业大学计算机学院  
Institution:  
国网浙江省电力有限公司信息通信分公司  
Institution:  
School of Computer Science,Northwestern Polytechnical University  
Institution:  
Information & Telecommunication Branch,State Grid Zhejiang Electric Power Co.,Ltd.

### 953. Robot Grasp Detection with Loss-Guided Collaborative Attention Mechanism

摘要: Grasp detection serves as the fundamental element for achieving successful grasping in robotic systems. The encoder-decoder structure has become widely adopted as the foundational architecture for grasp detection networks due to its inherent advantages of speed and accuracy. However, traditional network structures fail to effectively extract the essential features required for accurate grasping poses and neglect to eliminate the checkerboard artifacts caused by inversion convolution during decoding. Aiming at overcoming these challenges, we propose a novel generative grasp detection network (LGAR-Net2). A transposed convolution layer is employed to replace the bilinear interpolation layer in the decoder to remove the issue of uneven overlapping and consequently eliminate checkerboard artifacts. In addition, a loss-guided collaborative attention block (LGCA), which combines attention blocks with spatial pyramid blocks to enhance the attention to important regions of the image, is constructed to enhance the accuracy of information extraction. Validated on the Cornell public dataset using RGB images as the input, LGAR-Net2 achieves an accuracy of 97.7%, an improvement of 1.1% over the baseline network, and processes a single RGB image in just 15 ms.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001254702200001  
Indexed Date:  
2024-07-06  
Institution Address:  
Northwest Minzu Univ, Coll Elect Engn, Lanzhou 730030, Gansu, Peoples R China  
Institution Address:  
Gansu Engn Res Ctr Ecoenvironm Intelligent Network, Lanzhou 730030, Peoples R China  
Institution:  
Institution:  
Gansu Engn Res Ctr Ecoenvironm Intelligent Network

### 954. MFF-Net: Multiscale feature fusion semantic segmentation network for

摘要: BackgroundIn robot-assisted surgery, automatic segmentation of surgical instrument images is crucial for surgical safety. The proposed method addresses challenges in the craniotomy environment, such as occlusion and illumination, through an efficient surgical instrument segmentation network.MethodsThe network uses YOLOv8 as the target detection framework and integrates a semantic segmentation head to achieve detection and segmentation capabilities. A concatenation of multi-channel feature maps is designed to enhance model generalisation by fusing deep and shallow features. The innovative GBC2f module ensures the lightweight of the network and the ability to capture global information.ResultsExperimental validation of the intracranial glioma surgical instrument dataset shows excellent performance: 94.9% MPA score, 89.9% MIoU value, and 126.6 FPS.ConclusionsAccording to the experimental results, the segmentation model proposed in this study has significant advantages over other state-of-the-art models. This provides a valuable reference for the further development of intelligent surgical robots.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1478-5951  
eISSN:  
1478-596X  
入藏号:  
WOS:001095087700001  
PubMed ID:  
37932905  
Indexed Date:  
2023-11-16  
Institution Address:  
Tianjin Univ Technol, Sch Mech Engn, Tianjin Key Lab Adv Mechatron Syst Design & Intell, Tianjin, Peoples R China  
Institution Address:  
Tianjin Univ Technol, Natl Demonstrat Ctr Expt Mech & Elect Engn Educ, Tianjin, Peoples R China  
Institution Address:  
Tianjin Univ Technol, 391 Binshui West Rd, Tianjin 300384, Peoples R China  
Institution:  
Institution:  
Institution:

### 955. Lightweight Automatic Identification and Location Detection Model of

摘要: Automatic identification and location of farmland pests are an important direction of target detection research. The wide variety of pests and the similarity between pest categories make the automatic identification of farmland pests have some problems, such as high error rate and difficult identification. In order to achieve a better target for automatic identification and location of farmland pests, this paper proposes a lightweight pest detection model, and the network is the EfficientNet proposed by Google, which achieves the detection of 26 pests, the idea based on the classical Yolo target detection algorithm. First of all, features were extracted through the lightweight backbone, and then multiscale feature fusion is performed by PANet; finally, three feature matrices with different sizes were output to predict pests of different sizes. Using CIOU as the loss function of regression prediction better reflects the relative position of the prior box and the real box. The experimental results are compared with other lightweight algorithms, and the results show that the accuracy rate of the algorithm for identification and localization of agricultural pest in this paper is the highest and could reach 93.73%. Moreover, the model is lightweight and can be deployed on low-cost equipment, which reduces the cost of equipment and accurately predicts the status of pests and diseases in farmland. In practice, it is shown that the algorithm can effectively solve the problems of large number of pests, pest accumulation, background interference, and has strong robustness.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
SciELO Citation Index 中的 "被引频次":  
1  
被引频次合计:  
8  
ISSN:  
1530-8669  
eISSN:  
1530-8677  
入藏号:  
WOS:000730664800003  
Indexed Date:  
2021-12-24  
Institution Address:  
Univ Jinan, Jinan 250000, Peoples R China  
Institution Address:  
Plant Protect Stn, Jinan 250000, Peoples R China  
Institution:  
Institution:  
Plant Protect Stn

### 956. Traffic Sign Recognition Based on Semantic Scene Understanding and

摘要: Traffic sign recognition (TSR) plays an important role in driving assistance system and traffic safety insurance. However, existing methods focus on extracting features of traffic signs and ignore the constraints of spatial positional relationships between traffic signs and other objects in the scene. This way results in incorrectly detecting other similar objects as traffic signs and failing to detect very small traffic signs. A TSR method based on semantic scene understanding and structural traffic sign location is proposed in this study to solve the aforementioned problems. A scene structure model based on the constraints of spatial positional relationships between traffic signs and other objects is proposed to establish trusted search regions. An improved Light-weight RefineNet is used to analyze and understand a scene semantically and accurately and then segment objects in complicated environments precisely. A new network multiscale densely connected object detector (MDCOD) based on densely connected style, multiscale feature fusion, and improved K-means++ algorithms is proposed to recognize very small traffic signs. The trusted traffic signs are found by filtering false candidates outside the scene structure model. The proposed method is tested on Tsinghua-Tencent 100K and German Traffic Sign Detection Benchmark datasets and achieves accuracies of 92.8% and 99.90%, respectively, outperforming the existing methods.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
42  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
45  
ISSN:  
1524-9050  
eISSN:  
1558-0016  
入藏号:  
WOS:000751492300001  
Indexed Date:  
2022-01-31  
Institution Address:  
Nanchang Univ, Sch Software, Nanchang 330047, Jiangxi, Peoples R China  
Institution Address:  
Nanchang Univ, Sch Informat Engn, Nanchang 330031, Jiangxi, Peoples R China  
Institution Address:  
East China Normal Univ, Sch Comp Sci & Software Engn, Shanghai 200062, Peoples R China  
Institution:  
Institution:  
Institution:

### 957. Multiscale Parallel Algorithm for Early Detection of Tomato Gray Mold in

摘要: Plant disease detection technology is an important part of the intelligent agricultural Internet of Things monitoring system. The real natural environment requires the plant disease detection system to have extremely high real time detection and accuracy. The lightweight network MobileNetv2-YOLOv3 model can meet the real-time detection, but the accuracy is not enough to meet the actual needs. This study proposed a multiscale parallel algorithm MP-YOLOv3 based on the MobileNetv2-YOLOv3 model. The proposed method put forward a multiscale feature fusion method, and an efficient channel attention mechanism was introduced into the detection layer of the network to achieve feature enhancement. The parallel detection algorithm was used to effectively improve the detection performance of multiscale tomato gray mold lesions while ensuring the real-time performance of the algorithm. The experimental results show that the proposed algorithm can accurately and real-time detect multiscale tomato gray mold lesions in a real natural environment. The F1 score and the average precision reached 95.6 and 93.4% on the self-built tomato gray mold detection dataset. The model size was only 16.9 MB, and the detection time of each image was 0.022 s.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
5  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
9  
被引频次合计:  
10  
ISSN:  
1664-462X  
入藏号:  
WOS:000653902300001  
PubMed ID:  
34046045  
Indexed Date:  
2021-06-01  
Institution Address:  
Weifang Univ Sci & Technol, Shandong Prov Univ Lab Protected Hort, Blockchain Lab Agr Vegetables, Weifang, Peoples R China  
Institution:

### 958. UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for

摘要: Unmanned aerial vehicle (UAV) object detection plays a crucial role in civil, commercial, and military domains. However, the high proportion of small objects in UAV images and the limited platform resources lead to the low accuracy of most of the existing detection models embedded in UAVs, and it is difficult to strike a good balance between detection performance and resource consumption. To alleviate the above problems, we optimize YOLOv8 and propose an object detection model based on UAV aerial photography scenarios, called UAV-YOLOv8. Firstly, Wise-IoU (WIoU) v3 is used as a bounding box regression loss, and a wise gradient allocation strategy makes the model focus more on common-quality samples, thus improving the localization ability of the model. Secondly, an attention mechanism called BiFormer is introduced to optimize the backbone network, which improves the model's attention to critical information. Finally, we design a feature processing module named Focal FasterNet block (FFNB) and propose two new detection scales based on this module, which makes the shallow features and deep features fully integrated. The proposed multiscale feature fusion network substantially increased the detection performance of the model and reduces the missed detection rate of small objects. The experimental results show that our model has fewer parameters compared to the baseline model and has a mean detection accuracy higher than the baseline model by 7.7%. Compared with other mainstream models, the overall performance of our model is much better. The proposed method effectively improves the ability to detect small objects. There is room to optimize the detection effectiveness of our model for small and feature-less objects (such as bicycle-type vehicles), as we will address in subsequent research.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
28  
Web of Science 核心合集中的 "被引频次":  
250  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
18  
被引频次合计:  
270  
eISSN:  
1424-8220  
入藏号:  
WOS:001055769100001  
PubMed ID:  
37631727  
Indexed Date:  
2023-09-07  
Institution Address:  
Wuhan Inst Technol, Sch Elect & Informat Engn, Hubei Key Lab Opt Informat & Pattern Recognit, Wuhan 430205, Peoples R China  
Institution:

### 959. Multiscale Low-Light Image Enhancement Algorithm with Brightness

摘要: To address issues such as detail loss, artifacts, and unnatural appearance associated with current low-illumination image enhancement algorithms, a multiscale low-illumination image enhancement algorithm based on brightness equalization and edge enhancement is proposed in this study. Initially, an improved Sobel operator is employed to extract edge details, yielding an image with enhanced edge details. Subsequently, the brightness component (V) of the HSV color space is enhanced using Retinex, and brightness equalization is accomplished via improved Gamma correction, yielding an image with balanced brightness. The Laplacian weight graph, significance weight graph, and saturation weight graph are computed for the edge detail-enhanced image and brightness-balanced image, culminating in the generation of a normalized weight graph. This graph is then decomposed into a Gaussian pyramid, while the edge detail-enhanced image and brightness-balanced image are decomposed into a Laplacian pyramid. Finally, a multiscale pyramid fusion strategy is employed to merge the images, resulting in the final enhanced image. Experimental results demonstrate that the proposed algorithm outperforms existing algorithms on the LOL dataset in terms of average peak signal to noise ratio, structural similarity, and naturalness image quality evaluator. This algorithm effectively enhances the contrast and clarity of low-illumination images, resulting in images with richer detail information, improved color saturation, and considerably enhanced quality.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001284198700045  
Indexed Date:  
2024-09-18  
Institution Address:  
Liaoning Tech Univ, Sch Software, Huludao 125105, Liaoning, Peoples R China  
Institution Address:  
Liaoning Tech Univ, Dept Basic Educ, Huludao 125105, Liaoning, Peoples R China  
Institution Address:  
China Coal Technol & Engn Grp, Shenyang Res Inst, Fushun 113122, Liaoning, Peoples R China  
Institution Address:  
State Key Lab Coal Mine Safety Technol, Fushun 113122, Liaoning, Peoples R China  
Institution:  
Institution:  
Institution:  
China Coal Technol & Engn Grp  
Institution:  
State Key Lab Coal Mine Safety Technol

### 960. Few-Shot Learning Method for Multi-Scale Feature Aggregation

摘要: For most few-shot learning, there are problems such as insufficient feature information extraction in feature extraction, difficulty in accurately extracting important feature information in samples, and the diversity of samples within a class may lead to the deviation of class center points. A few-shot learning method(MSFA) for multi-scale feature aggregation is proposed. Specifically, the method uses a multi-scale generation module to generate feature information of multiple different scales about all training samples, secondly self-attention is used to aggregate important feature information of different scales, and important feature information of different scales are spliced. This achieves a more accurate feature representation about the image. Finally, the distance between each query set sample and the class prototype and the average distance from each sample in the class are calculated separately, and the final distance is obtained in a weighted manner. A large number of experiments are carried out on the three datasets of miniImageNet, tieredImageNet and Standford Dogs. The experimental results show that the poposed method can greatly improve the classification performance of the baseline method, especially on the miniImageNet dataset, in 5-way 1-shot and 5-way 5-shot setting, compared with the Prototypical Network method, the classification accuracy is improved by 7.42 and 6.28 percentage points, respectively.  
摘要:  
针对大多数小样本学习在特征提取中存在特性信息提取不足、难以准确地提取样本中的重要特征信息以及类内样本多样性可能导致类中心点偏离等问题。提出一种多尺度特征聚合的小样本学习方法(MSFA)。具体来说,该方法利用多尺度生成模块生成关于全部训练样本的多种不同尺度的特征信息,使用自注意力聚合不同尺度的重要特征信息,并将不同尺度的重要特征信息进行拼接,以此来实现关于图像更为准确的特征表达。分别计算每个查询集样本与类原型的距离以及与类内各样本间距离的平均值,并以加权方式得出最终距离。在miniImageNet、tiered-ImageNet和Standford Dogs三个数据集上进行大量的实验,实验结果表明:提出的方法可以大幅提升基线方法的分类性能,特别是在miniImageNet数据集上,在5-way 1-shot和5-way 5-shot设置中,相较于Prototypical Network方法,分类准确率分别提升7.42和6.28个百分点。  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7544082  
Indexed Date:  
2023-12-08  
Institution Address:  
福建工程学院计算机科学与数学学院, 福州, 福建 350118, 中国  
Institution Address:  
福建工程学院计算机科学与数学学院;;福建省大数据挖掘与应用重点实验室, ;;福建省大数据挖掘与应用重点实验室, 福州;;福州, ;; 350118;;350118  
Institution Address:  
College of Computer Science and Mathematics,Fujian University of Technology, Fuzhou, Fujian 350118, China  
Institution Address:  
College of Computer Science and Mathematics,Fujian University of Technology;;Fujian Key Laboratory of Big Data Mining and Applications, ;;Fujian Key Laboratory of Big Data Mining and Applications, Fuzhou;;Fuzhou, ;; 350118;;350118  
Institution:  
福建工程学院计算机科学与数学学院  
Institution:  
福建工程学院计算机科学与数学学院;;福建省大数据挖掘与应用重点实验室  
Institution:  
College of Computer Science and Mathematics,Fujian University of Technology  
Institution:  
College of Computer Science and Mathematics,Fujian University of Technology;;Fujian Key Laboratory of Big Data Mining and Applications

### 961. Co-learning-assisted progressive dense fusion network for cardiovascular

摘要: Electrocardiograms (ECGs) and phonocardiograms (PCGs) are two modalities to provide complementary diagnostic information for improving the early detection accuracy of cardiovascular diseases (CVDs). Existing multi-modality methods mainly used the early or late feature fusion strategy, which did not simultaneously utilize the complementary information contained in low-level detail features and high-level semantic features of different modalities. Meanwhile, they were specially designed for the multi-modality scenario with both ECGs and PCGs, without considering the missing-modality scenarios with only ECGs or PCGs in clinical practice. To address these challenges, we developed a Co-learning-assisted Progressive Dense fusion network (CPDNet) for end-to-end CVD detection, with a three-branch interweaving architecture consisting of ECG and PCG modality-specific encoders and a progressive dense fusion encoder, which could be used for both multi-modality and missing-modality scenarios. Specifically, we designed a novel progressive dense fusion strategy, which not only progressively fused multi-level complementary information of different modalities from low-level details to high-level se -mantics, but also employed the dense fusion during feature fusion at each level to further enrich available multi-modality information through mutual guidance of features at different levels. Meanwhile, the strategy integrated cross-modality region-aware and multi-scale feature optimization modules to fully evaluate the contributions of different modalities and signal regions and enhance the feature extraction ability of the network for multi-scale target regions. Moreover, we designed a novel co-learning strategy to guide the learning process of the CPDNet by combining intra-modality and joint losses, which made each encoder well-trained. This strategy could not only assist our fusion strategy by making modality-specific encoders provide sufficiently discriminative features for the fusion encoder, but also enable the CPDNet to robustly handle missing-modality scenarios by independently using the corresponding modality-specific encoder. Experimental results on public and private datasets demonstrated that our method not only outperformed state-of-the-art multi-modality methods by at least 5.05% for average accuracy in the multi-modality scenario, but also achieved better performance than single-modality models in the missing-modality scenarios.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001105408200001  
Indexed Date:  
2023-12-17  
Institution Address:  
Huazhong Univ Sci & Technol, Britton Chance Ctr Biomed Photon, Wuhan Natl Lab Optoelect, Wuhan, Hubei, Peoples R China  
Institution Address:  
Huazhong Univ Sci & Technol, Collaborat Innovat Ctr Biomed Engn, Sch Engn Sci, MoE Key Lab Biomed Photon, Wuhan, Hubei, Peoples R China  
Institution Address:  
Huazhong Univ Sci & Technol, Tongji Hosp, Tongji Med Coll, Div Cardiol,Dept Internal Med, Wuhan, Hubei, Peoples R China  
Institution:  
Institution:  
Institution:

### 962. Target Detection of Remote-Sensing Images Based on Improved YOLOv5

摘要: Although target detection technology has advanced,many challenges still exist in the detection of remote-sensing images.An improved YOLOv5-based remote-sensing image target detection algorithm is proposed to address the issues of low target detection accuracy caused by complex backgrounds,large target scale differences,and arbitrary target orientation in remote-sensing images.First,a joint multiscale feature enhancement network with attention is constructed to fully fuse high-level and low-level features such that the feature layers contain semantic and rich detailed information.During the fusion process,the designed feature focusing module is used to help the model select key features and suppress irrelevant information.Second,a Receptive Field Block(RFB) is used to update the fused feature map and expand the receptive field of the feature map to reduce feature information loss.Finally,by adding rotation angles to the targets and using circular smooth labels to transform the regression problem into a classification problem,the accuracy of remote-sensing target localization is improved.The experimental results on the a large-scale Dataset for Object deTection in Aerial images(DOTA) show that compared with the YOLOv5 algorithm,the mean Average Precision(mAP) when the Intersection over Union(IoU) values of the proposed algorithm are 0.5 and 0.5-0.95(mAP@0.5 and mAP@0.5∶0.95) increase by 7.3 and 3.3 percentage points,respectively.This can significantly improve the detection accuracy of remote-sensing image targets in a complex background and improve the missing and false detection of remote-sensing targets.  
摘要:  
目前目标检测技术虽然已经趋于成熟,但是对遥感图像的检测仍存在不少挑战。针对遥感图像的背景复杂、目标尺度差异大、目标方向任意等特点造成目标检测精度低下的问题,提出一种基于改进YOLOv5的遥感图像目标检测算法。首先,构建一种联合注意力的多尺度特征增强网络,充分融合高低层特征,使特征层具有语义信息的同时包含丰富的细节信息,并在融合过程中利用设计的特征聚焦模块帮助模型选择关键特征,抑制无关信息。其次,使用感受野模块(RFB)对融合后的特征图进行更新,扩大特征图的感受野,减少特征信息损失。最后,对目标增加旋转角度,并采用圆形平滑标签将回归问题转化成分类问题,提高遥感目标定位的准确性。在用于航拍图像目标检测的大规模数据集(DOTA)上的实验结果表明,与YOLOv5算法相比,所提算法的交并比(IoU)为0.5和0.5~0.95时的平均精度均值(mAP@0.5和mAP@0.5∶ 0.95)分别提高了7.3和3.3个百分点,能够明显提高复杂背景下遥感图像目标的检测精度,并改善对遥感目标的漏检和误检情况。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
被引频次合计:  
3  
ISSN:  
1000-3428  
入藏号:  
CSCD:7698346  
Indexed Date:  
2024-06-29  
Institution Address:  
辽宁工程技术大学软件学院, 葫芦岛, 辽宁 125105, 中国  
Institution Address:  
College of Software,Liaoning Technical University, Huludao, Liaoning 125105, China  
Institution:  
辽宁工程技术大学软件学院  
Institution:  
College of Software,Liaoning Technical University

### 963. Target Tracking Algorithm Based on Siamese Network of Feature

摘要: In the target tracking sequences, it is difficult to identify the target because of the complex background and large-scale changes of the target. To solve this problem, a target tracking algorithm based on feature optimization model in the Siamese network is proposed. First, the deep network is constructed to extract the deep semantic information effectively. Then, the hourglass network is used to encode the global features of the multi-scale feature map, and the encoded features are normalized to obtain the effective target features. Finally, a feature optimization model is constructed, and the features obtained by decoding are used as selectors to identify and enhance the effective features of the original feature map. In order to further improve the generalization ability of the model, the attention mechanism is introduced to adaptively weigh the target features to adapt to the scene changes. The proposed algorithm is tested on two standard tracking data sets including OTB100 and VOT2018. The success rate in the OTB100 is 0.648, the prediction accuracy is 0.853, and the real-time performance is 59.5 frame/s; the test accuracy in the VOT2018 is 0.536, the expected average coverage rate is 0.192, and the real-time performance is 44.3 frame/s. The test results prove the effectiveness of the proposed algorithm.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:000823188100026  
Indexed Date:  
2022-08-13  
Institution Address:  
Inner Mongolia Univ Sci & Technol, Sch Informat Engn, Baotou 014010, Inner Mongolia, Peoples R China  
Institution Address:  
Mongolia Ind Univ, Sch Informat Engn, Hohhot 010051, Inner Mongolia, Peoples R China  
Institution Address:  
Inner Mongolia Key Lab Patten Recognit & Intellig, Baotou 014010, Inner Mongolia, Peoples R China  
Institution:  
Institution:  
Mongolia Ind Univ  
Institution:  
Inner Mongolia Key Lab Patten Recognit & Intellig

### 964. Multi-scale sampling attention graph convolutional networks for

摘要: Skeleton-based action recognition has attracted increasing interest in recent years. With the flexibility of modeling long-range dependency of joints, the self-attention module has served as the basic component in skeleton-based action recognition. However, the global receptive field of self-attention is not conducive to the modeling of skeleton locality, and the self-attention model is imbued with less inductive bias, which leads to overfitting. In this paper, we propose an attention graph convolutional network (AGCN) with multi- scale sampling to effectively model the local and global features of the skeleton. Firstly, we propose two extreme sampling strategies for generating and ordering neighboring nodes of root nodes. A local-first sampling method is introduced to construct local graph windows, and a global-first sampling method is proposed to assemble long-range joints for constructing global graph windows. The local-first sampling and global-first sampling introduce more skeleton-specific inductive biases to regularize the model capacity. Secondly, the AGCN combines the self-attention mechanism with graph convolution operation, which alleviates the over- smoothing of graph convolution and preserves the translation invariant. Based on the multi-scale sampling strategy, the AGCN can effectively model the locality and non-locality of the skeleton. Finally, by coupling the aforementioned proposals, we develop a two-pathway model for multi-scale feature fusion. Extensive experiments demonstrate that our model could achieve comparable performance with state-of-the-art works on the NTU RGB+D 60, NTU RGB+D 120, the UAV-HUMAN and NW-UCLA datasets.  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:001283932400001  
Indexed Date:  
2024-08-14  
Institution Address:  
Shandong Univ, Ctr Robot, Sch Control Sci & Engn, Jinan 250061, Peoples R China  
Institution Address:  
Univ Calif Los Angeles, Dept Elect & Comp Engn, Los Angeles, CA 90095 USA  
Institution:  
Institution:

### 965. LCCDMamba: Visual State Space Model for Land Cover Change Detection of

摘要: Land cover change detection (LCCD) is a crucial research topic for rational planning of land use and facilitation of sustainable land resource growth. However, due to the complexity of LCCD tasks, integrating global and local features and fusing contextual information from remote sensing features are essential. Recently, with the advent of Mamba, which maintains linear time complexity and high efficiency in processing long-range data, it offers a new solution to address feature-fusion challenges in LCCD. Therefore, a novel visual state space model (SSM) for Land Cover Change Detection (LCCDMamba) is proposed, which uses Siam-VMamba as a backbone to extract multidimensional land cover features. To fuse the change information across difference temporal, multiscale information spatio-temporal fusion (MISF) module is designed to aggregate difference information from bitemporal features. The proposed MISF comprises multi-scale feature aggregation (MSFA), which utilizes strip convolution to aggregate multiscale local change information of bitemporal land cover features, and residual with SS2D (RSS) which employs residual structure with SS2D to capture global feature differences of bitemporal land cover features. To enhance the correlation of change features across different dimensions, in the decoder, we design a dual token modeling SSM (DTMS) through two token modeling approaches. This preserves high-dimensional semantic features and thus ensures that the multiscale change information across various dimensions will not be lost during feature restoration. Experiments have been conducted on WHU-CD, LEVIR-CD, and GVLM datasets, and the results demonstrate that LCCDMamba achieves F1 scores of 94.18%, 91.68%, and 87.14%, respectively, outperforming all the models compared.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001432389100002  
Indexed Date:  
2025-03-06  
Institution Address:  
Macao Polytech Univ, Fac Appl Sci, Taipa 999078, Macau, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Collaborat Innovat Ctr Atmospher Environm & Equipm, Nanjing 210044, Peoples R China  
Institution:  
Institution:

### 966. Multi-scale feature fusion for single image novel view synthesis

摘要: Single image novel view synthesis allows the generation of target images with different views from a single input image. Pixel generation methods are one of the main approaches for novel view synthesis, with previous methods typically using the input image to infer the target image in the new view. However, only features from input images in the source view might not be sufficient to generate a good target image, especially when only a single input image is available. In this paper, we fuse features from an input and a warped image to collaboratively generate pixels in the new view, with the warped image as an intermediate output generated by projecting pixels of the input image onto the target view via an estimated depth. Since the estimated depth and the generated warped image are not perfect, errors will be introduced when generating target pixels. To alleviate these and to ensure better channel information between the features from input and warped image, channel attention blocks are employed. In addition, in order to use skip connections for better novel view synthesis results, encoder features in different layers from the input image are transformed to the target view via multi-resolution depths. Here, instead of downsampling a single full-resolution depth to several lower-resolution depths, we adopt a multi-scale depth estimation network to predict multiple depths at different resolutions. Experimental results on benchmark datasets show that our method gives excellent view synthesis results and outperforms other state-of-the-art novel view synthesis methods.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:001260913500001  
Indexed Date:  
2024-07-10  
Institution Address:  
Loughborough Univ, Dept Comp Sci, Loughborough, England  
Institution:

### 967. HRCUNet: Hierarchical Region Contrastive Learning for Segmentation of

摘要: Segmenting breast tumors from dynamic contrast-enhanced magnetic resonance images is a critical step in the early detection and diagnosis of breast cancer. However, this task becomes significantly more challenging due to the diverse shapes and sizes of tumors, which make it difficult to establish a unified perception field for modeling them. Moreover, tumor regions are often subtle or imperceptible during early detection, exacerbating the issue of extreme class imbalance. This imbalance can lead to biased training and challenge accurately segmenting tumor regions from the predominant normal tissues. To address these issues, we propose a hierarchical region contrastive learning approach for breast tumor segmentation. Our approach introduces a novel hierarchical region contrastive learning loss function that addresses the class imbalance problem. This loss function encourages the model to create a clear separation between feature embeddings by maximizing the inter-class margin and minimizing the intra-class distance across different levels of the feature space. In addition, we design a novel Attention-based 3D Multi-scale Feature Fusion Residual Module to explore more granular multi-scale representations to improve the feature learning ability of tumors. Extensive experiments on two breast DCE-MRI datasets demonstrate that the proposed algorithm is more competitive against several state-of-the-art approaches under different segmentation metrics.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1532-0626  
eISSN:  
1532-0634  
入藏号:  
WOS:001354204600001  
Indexed Date:  
2024-11-21  
Institution Address:  
Xiamen Univ, Inst Artificial Intelligence, Xiamen, Peoples R China  
Institution Address:  
Xiamen Univ, Dept Artificial Intelligence, Xiamen, Peoples R China  
Institution Address:  
Stanford Univ, Dept Psychiat & Behav Sci, Stanford, CA USA  
Institution Address:  
Xiamen Univ, Xiangan Hosp Xiamen Univ, Sch Med, Fujian Key Lab Precis Diag & Treatment Breast Canc, Xiamen 361100, Peoples R China  
Institution Address:  
Xiamen Univ, Natl Inst Data Sci Hlth & Med, Xiamen, Peoples R China  
Institution Address:  
Xiamen Univ, Xiangan Hosp Xiamen Univ, Xiamen Res Ctr Clin Med Breast & Thyroid Canc, Sch Med, Xiamen 361100, Peoples R China  
Institution Address:  
Xiamen Univ, Xiangan Hosp, Canc Ctr, Sch Med, Xiamen, Peoples R China  
Institution Address:  
Xiamen Univ, Xiangan Hosp, Sch Med, Dept Breast & Thyroid Surg, Xiamen, Peoples R China  
Institution Address:  
Xiamen Univ, Xiangan Hosp Xiamen Univ, Sch Med, Xiamen Key Lab Endocrine Related Canc Precis Med, Xiamen 361100, Peoples R China  
Institution Address:  
Third Affiliated Hosp Kunming Med Univ, Yunnan Canc Hosp, Peking Univ Yunnan Canc Hosp, Kunming, Yunnan, Peoples R China  
Institution Address:  
Peking Univ, Canc Hosp, Kunming, Yunnan, Peoples R China  
Institution Address:  
Xiamen Univ, Canc Res Ctr, Sch Med, Xiamen, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 968. U-Net-based Segmentation of Crack on Dam Surface Using Separable

摘要: Detection of defects on dam surface is a key part in safety inspection for water conservancy projects.However,dam surface images under complex environment are faced with large interference noise and unbalanced pixels,making it difficult to segment cracks on dam surface accurately.This paper proposes a U-Net-based method for crack segmentation using separable residual convolution and semantic compensation.At the coding end of U-Net,a larger separable residual convolution module is constructed to replace the conventional convolution module,so as to expand the receptive field of feature layer and avoid losing the crack boundary information.At the same time,at the decoding end,a semantic feature compensation module is added to improve the effect of multi-scale feature fusion.Take the focal loss function and center loss function as the objective function,the loss weights of crack foreground and difficult samples are increased to improve the accuracy of classification.Experimental results on a self-made dam surface crack dataset of a hydropower station in Southwest China show that the F1 value and Intersection over Union(IoU)of the proposed method reach 69.89% and 53.72% respectively.The segmentation effect of this method is better than that of the traditional methods such as SegNet and FCN-8S,and it has better performance in small crack recognition.  
摘要:  
坝面缺陷检测是水利枢纽安全巡检的关键环节,但复杂环境下坝面图像存在干扰噪声大和像素不均衡等问题,造成坝面裂缝难以精细分割。提出一种利用可分离残差卷积和语义补偿的U-Net裂缝分割方法。在U-Net网络的编码端构建更大尺寸的可分离残差卷积模块替换常规卷积模块,从而扩大特征层感受野并避免丢失裂缝边界信息,同时在解码端增加语义特征补偿模块改善多尺度特征融合效果,将焦点损失函数和中心损失函数作为目标函数,加大裂缝前景与困难样本的损失权重以提高分类准确度。在自制西南某水电站坝面裂缝数据集上的实验结果表明,该方法的F1值和交并比分别达到69.89%与53.72%,分割效果较SegNet、FCN-8S等传统方法更优,对细小裂缝区域的识别能力更强。  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1000-3428  
入藏号:  
CSCD:6949962  
Indexed Date:  
2021-07-15  
Institution Address:  
特殊环境机器人技术四川省重点实验室;;西南科技大学信息工程学院, 特殊环境机器人技术四川省重点实验室;;, 绵阳;;绵阳, 四川;;四川 621000;;621000, 中国  
Institution Address:  
特殊环境机器人技术四川省重点实验室;;西南科技大学信息工程学院;;清华四川能源互联网研究院, 特殊环境机器人技术四川省重点实验室;;;;, 绵阳;;绵阳;;成都, 四川;;四川;; 621000;;621000;;610213, 中国  
Institution Address:  
Special Environment Robot Technology Key Laboratory of Sichuan Province;;School of Information Engineering,Southwest University of Science and Technology, Special Environment Robot Technology Key Laboratory of Sichuan Province;;, Mianyang;;Mianyang, Sichuan;;Sichuan 621000;;621000  
Institution Address:  
Special Environment Robot Technology Key Laboratory of Sichuan Province;;School of Information Engineering,Southwest University of Science and Technology;;Tsinghua Sichuan Energy Internet Research Institute, Special Environment Robot Technology Key Laboratory of Sichuan Province;;;;, Mianyang;;Mianyang;;Chengdu, Sichuan;;Sichuan;; 621000;;621000;;610213  
Institution:  
特殊环境机器人技术四川省重点实验室;;西南科技大学信息工程学院  
Institution:  
特殊环境机器人技术四川省重点实验室;;西南科技大学信息工程学院;;清华四川能源互联网研究院  
Institution:  
Special Environment Robot Technology Key Laboratory of Sichuan Province;;School of Information Engineering,Southwest University of Science and Technology  
Institution:  
Special Environment Robot Technology Key Laboratory of Sichuan Province;;School of Information Engineering,Southwest University of Science and Technology;;Tsinghua Sichuan Energy Internet Research Institute

### 969. The bi-long short-term memory based on multiscale and mesoscale feature

摘要: Accurate power load prediction is beneficial to the efficient use of electric energy and the orderly development of power systems. Given the strong volatility and complexity of power load series, a hybrid load forecasting method based on multiscale and mesoscale information fusion, signal decomposition, model optimization, and bi-longshort-term memory (BiLSTM) is proposed. Firstly, the load sequence is analyzed on different time scales, and the extracted multi-scale information and mesoscale information are fused to improve the perception ability. Secondly, the empirical wavelet transform (EWT) with adaptive decomposition ability is used to decompose the sequence and extract the rich feature information. Thirdly, the complexity, volatility, and uncertainty of each mode component were analyzed, the data features were fully mined, and the feature fusion was carried out by the TOPISIS evaluation method. The BiLSTM model and the GWO-BiLSTM model are used to predict the low- frequency component and the high-frequency component, respectively. The optimization of Grey Wolf optimization (GWO) algorithm can improve the BiLSTM model's ability to learn long-term time series. Finally, the analysis of application examples shows that compared with various prediction models, the prediction error of mixed model EWT-SGEO-BiLSTM is the smallest, MAPE is as low as 1.07 %, and goodness of fit R2 2 is 0.99 which verifies the accuracy and applicability of the intelligent model.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
ISSN:  
1568-4946  
eISSN:  
1872-9681  
入藏号:  
WOS:001339164100001  
Indexed Date:  
2024-10-31  
Institution Address:  
Ping Ding Shan Univ, Sch Math & Stat, Ping Ding Shan 467000, Henan, Peoples R China  
Institution Address:  
Chihlee Univ Sci, Dept Informat Management, New Taipei 220305, Taiwan  
Institution Address:  
Asia Eastern Univ Sci & Technol, Dept Informat Management, New Taipei 22064, Taiwan  
Institution Address:  
Yuan Ze Univ, Dept Informat Management, Taoyuan City 320315, Taiwan  
Institution:  
Ping Ding Shan Univ  
Institution:  
Chihlee Univ Sci  
Institution:  
Asia Eastern Univ Sci & Technol  
Institution:

### 970. DRSI-Net: Dual-residual spatial interaction network for multi-person

摘要: Multi -person pose estimation (MPPE), which aims to locate the key points for all persons in the frames, is an active research branch of computer vision. Variable human poses and complex scenes make MPPE dependent on local details and global structures; their absence may cause key point feature misalignment. In this case, highorder spatial interactions that can effectively link the local and global information of features are particularly important. However, most methods do not include spatial interactions. A few methods have low-order spatial interactions, but achieving a good balance between accuracy and complexity is challenging. To address the above problems, a dual-residual spatial interaction network (DRSI-Net) for MPPE with high accuracy and low complexity is proposed herein. Compared to other methods, DRSI-Net recursively performs residual spatial information interactions on the neighbouring features so that more useful spatial information can be retained and more similarities can be obtained between shallow and deep extracted features. The channel and spatial dual attention mechanism introduced in the multi-scale feature fusion also helps the network to adaptively focus on features relevant to the target key points and further refine the generated poses. Simultaneously, by optimising the interactive channel dimensions and dividing the gradient flow, the spatial interaction module is designed to be lightweight, thus reducing the complexity of the network. According to the experimental results on the COCO dataset, the proposed DRSI-Net outperforms other state -of -the -art methods in accuracy and complexity.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0950-7051  
eISSN:  
1872-7409  
入藏号:  
WOS:001236505200001  
Indexed Date:  
2024-06-06  
Institution Address:  
Shanghai Univ, Sch Commun & Informat Engn, Shanghai, Peoples R China  
Institution:

### 971. Optimized YOLOv7 for Small Target Detection in Aerial Images Captured by

摘要: It is challenging to detect small targets in aerial images captured by drones due to variations in target sizes and occlusions arising from the surrounding environment. This study proposes an optimized object detection algorithm based on YOLOv7 to address the above-mentioned challenges. The proposed method comprises the design of a Genetic Kmeans (1IoU) clustering algorithm to obtain customized anchor boxes that more significantly apply to the dataset. Moreover, the SPPFCSPC\_group structure is optimized using group convolutions to reduce model parameters. The fusion of Spatial Pyramid Pooling-Fast (SPPF) and Cross Stage Partial (CSP) structures leads to increased detection accuracy and enhanced multi-scale feature fusion network. Furthermore, a Detect Head is incorporated into the classification phase for more accurate position and class predictions. According to experimental findings, the optimized YOLOv7 algorithm performs quite well on the VisDrone2019 dataset in terms of detection accuracy. Compared with the original YOLOv7 algorithm, the optimized version shows a 0.18% increase in the Average Precision (AP), a reduction of 5.7 M model parameters, and a 1.12 Frames Per Second (FPS) improvement in the frame rate. With the above described enhancements in AP and parameter reduction, the precision of small target detection and the real-time detection speed are increased notably. In general, the optimized YOLOv7 algorithm offers superior accuracy and real-time capability, thus making it well-suited for small target detection tasks in real-time drone aerial photography.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2158-107X  
eISSN:  
2156-5570  
入藏号:  
WOS:001089171100001  
Indexed Date:  
2023-11-08  
Institution Address:  
Liaoning Petrochem Univ, Sch Informat & Control Engn, Fushun, Peoples R China  
Institution:

### 972. CoFiNet: Unveiling camouflaged objects with multi-scale finesse

摘要: Camouflaged Object Detection (COD) is a critical aspect of computer vision aimed at identifying concealed objects, with applications spanning military, industrial, medical and monitoring domains. To address the problem of poor detail segmentation effect, we introduce a novel method for camouflaged object detection, named CoFiNet. Our approach primarily focuses on multi-scale feature fusion and extraction, with special attention to the model's segmentation effectiveness for detailed features, enhancing its ability to effectively detect camouflaged objects. CoFiNet adopts a coarse-to-fine strategy. A multi-scale feature integration module is laveraged to enhance the model's capability of fusing context feature. A multi-activation selective kernel module is leveraged to grant the model the ability to autonomously alter its receptive field, enabling it to selectively choose an appropriate receptive field for camouflaged objects of different sizes. During mask generation, we employ the dual-mask strategy for image segmentation, separating the reconstruction of coarse and fine masks, which significantly enhances the model's learning capacity for details. Comprehensive experiments were conducted on four different datasets, demonstrating that CoFiNet achieves state-of-the-art performance across all datasets. The experiment results of CoFiNet underscore its effectiveness in camouflaged object detection and highlight its potential in various practical application scenarios.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:001349451800001  
Indexed Date:  
2024-11-14  
Institution Address:  
Univ Chinese Acad Sci, Sch Emergency Management Sci & Engn, 1 Yanqihu East Rd, Beijing 101400, Peoples R China  
Institution Address:  
Beijing Inst Technol, Southeast Acad Informat Technol, 1998 Licheng Middle Ave, Putian 351100, Peoples R China  
Institution Address:  
Beijing Inst Technol, Sch Comp Sci & Technol, 5 Zhongguancun South St, Beijing 100086, Peoples R China  
Institution:  
Institution:  
Institution:

### 973. Automatic Bird Species Recognition from Images with Feature Enhancement

摘要: Featured Application The research introduced in the paper develops an automatic bird species recognition system powered by feature enhancement and contrast learning, aimed at advancing ecological conservation and biological research. This system bolsters the precision of identifying bird species, aiding in the protection of endangered birds and automating the monitoring of bird populations and their migratory behaviors. It also supports in-depth behavioral and ecological research and evaluates the impact of human activities on avian life. The technology's potential applications are vast, including its use in citizen science initiatives, environmental impact assessments, educational programs, and the tourism sector, where it can provide real-time species identification, thereby enriching eco-tourism experiences and raising biodiversity awareness.Abstract Accurate bird species recognition is crucial for ecological conservation, wildlife monitoring, and biological research, yet it poses significant challenges due to the high variability within species and the subtle similarities between different species. This paper introduces an automatic bird species recognition method from images that leverages feature enhancement and contrast learning to address these challenges. Our method incorporates a multi-scale feature fusion module to comprehensively capture information from bird images across diverse scales and perspectives. Additionally, an attention feature enhancement module is integrated to address noise and occlusion within images, thus enhancing the model's robustness. Furthermore, employing a siamese network architecture allows effective learning of common features within instances of the same class and distinctions between different bird species. Evaluated on the CUB200-2011 dataset, our proposed method achieves state-of-the-art performance, surpassing existing methods with an accuracy of 91.3% and F1 score of 90.6%. Moreover, our approach showcases a notable advantage in scenarios with limited training data. When utilizing only 5% of the training data, our model still achieves a recognition accuracy of 65.2%, which is significantly higher than existing methods under similar data constraints. Notably, our model exhibits faster execution times compared to existing methods, rendering it suitable for real-time applications.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001232749100001  
Indexed Date:  
2024-06-03  
Institution Address:  
Beijing Forestry Univ, Sch Informat Sci & Technol, Beijing 100083, Peoples R China  
Institution Address:  
Natl Forestry & Grassland Adm, Engn Res Ctr Forestry Oriented Intelligent Informa, Beijing 100083, Peoples R China  
Institution:  
Institution:  
Natl Forestry & Grassland Adm

### 974. DSNet: Dual-stream multi-scale fusion network for low-quality 3D face

摘要: 3D face recognition (FR) has become increasingly widespread due to the illumination invariance and pose robustness of 3D face data. Most existing 3D FR methods can only achieve excellent performance on complete and smooth faces. However, low-quality 3D FR with missing facial features still suffers from insufficient discriminative feature extraction for visible face regions. In this paper, we propose a dual-stream multi-scale fusion network (DSNet) for low-quality 3D FR. First, in the first stream, we design a new multi-scale local and global feature fusion network, which consists of an enhanced shallow feature extraction module, an enhanced deep feature extraction module, and a layered multi-scale feature correlation fusion module, aiming to obtain more discriminative details and category information of the facial visible region, reducing the interference of similar features and the redundancy of the same features. Second, we also introduced a capsule network as the second stream to enhance the expression of 3D facial spatial position information, thereby further improving the performance of low-quality 3D FR with missing facial features. We conduct extensive experiments on low-quality datasets (Lock3DFace, KinectFaceDB, and IIIT-D) and cross-quality datasets synthesized by Bosphorus. These results show that our proposed DSNet can achieve state-of-the-art recognition performance and exhibit excellent performance on low-quality 3D faces with missing facial features.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
2  
eISSN:  
2158-3226  
入藏号:  
WOS:001044500100002  
Indexed Date:  
2023-08-26  
Institution Address:  
Beijing Univ Posts & Telecommun, Beijing 100876, Peoples R China  
Institution:

### 975. Efficient Small Object Detection You Only Look Once: A Small Object

摘要: Aerial images have distinct characteristics, such as varying target scales, complex backgrounds, severe occlusion, small targets, and dense distribution. As a result, object detection in aerial images faces challenges like difficulty in extracting small target information and poor integration of spatial and semantic data. Moreover, existing object detection algorithms have a large number of parameters, posing a challenge for deployment on drones with limited hardware resources. We propose an efficient small-object YOLO detection model (ESOD-YOLO) based on YOLOv8n for Unmanned Aerial Vehicle (UAV) object detection. Firstly, we propose that the Reparameterized Multi-scale Inverted Blocks (RepNIBMS) module is implemented to replace the C2f module of the Yolov8n backbone extraction network to enhance the information extraction capability of small objects. Secondly, a cross-level multi-scale feature fusion structure, wave feature pyramid network (WFPN), is designed to enhance the model's capacity to integrate spatial and semantic information. Meanwhile, a small-object detection head is incorporated to augment the model's ability to identify small objects. Finally, a tri-focal loss function is proposed to address the issue of imbalanced samples in aerial images in a straightforward and effective manner. In the VisDrone2019 test set, when the input size is uniformly 640 x 640 pixels, the parameters of ESOD-YOLO are 4.46 M, and the average mean accuracy of detection reaches 29.3%, which is 3.6% higher than the baseline method YOLOv8n. Compared with other detection methods, it also achieves higher detection accuracy with lower parameters.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
1424-8220  
入藏号:  
WOS:001351076200001  
PubMed ID:  
39517964  
Indexed Date:  
2024-11-16  
Institution Address:  
Wuhan Univ Technol, Sch Informat Engn, Wuhan 430070, Peoples R China  
Institution Address:  
Wuhan Chuyan Informat Technol Co Ltd, Wuhan 430030, Peoples R China  
Institution:  
Institution:  
Wuhan Chuyan Informat Technol Co Ltd

### 976. TransFusionNet: Semantic and Spatial Features Fusion Framework for Liver

摘要: Liver cancer is one of the most common malignant diseases worldwide. Segmentation and reconstruction of liver tumors and vessels in CT images can provide convenience for physicians in preoperative planning and surgical intervention. In this paper, we introduced a TransFusionNet framework, which consists of a semantic feature extraction module, a local spatial feature extraction module, an edge feature extraction module, and a multi-scale feature fusion module to achieve fine-grained segmentation of liver tumors and vessels. In addition, we applied the transfer learning approach to pre-train using public datasets and then fine-tune the model to further improve the fitting effect. Furthermore, we proposed an intelligent quantization scheme to compress the model weights and achieved high performance inference on JetsonTX2. The TransFusionNet framework achieved mean IoU of 0.854 in vessel segmentation task, and achieved mean IoU of 0.927 in liver tumor segmentation task. When profiling the Computational Performance of the quantized inference, our quantized model achieved 4TFLOPs on Node with NVIDIA RTX3090 and 132GFLOPs on JetsonTX2. This unprecedented segmentation effect solves the accuracy and performance bottleneck of automated segmentation to a certain extent.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
5  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
被引频次合计:  
14  
ISSN:  
2168-2194  
eISSN:  
2168-2208  
入藏号:  
WOS:000965360300001  
PubMed ID:  
36112546  
Indexed Date:  
2023-07-06  
Institution Address:  
China Univ Petr, Dept Comp Sci & Technol, Qingdao 266580, Shandong, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, High Performance Comp Res Ctr, Beijing 100190, Peoples R China  
Institution Address:  
Qingdao Univ Sci & Technol, Sch Math & Phys, Qingdao 266061, Shandong, Peoples R China  
Institution Address:  
Qingdao Municipal Hosp, Minimally Invas Intervent Therapy Ctr, Qingdao 266011, Shandong, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 977. Multi-scale multi-modal fusion for object detection in autonomous

摘要: Fusion object detection using camera and LiDAR information in autonomous driving is still a challenging task, the difference between sensor data increases the difficulty of data fusion. To address this issue, we propose a multi-scale selective kernel fusion(MSSKF) method and demonstrate its practical utility by using LiDAR-camera fusion in object detection network. Specifically, a multi-scale feature fusion module that uses multi-scale convolution to separate the feature expression of multi-modal information and calculates the weight of each modal feature channel is proposed. We use the idea of multi-scale convolution and selection kernel to complete multi-modal fusion in object detection, which is conducive to solving the problem that the image and point cloud fusion are difficult to match due to the difference in data structure, and the complementarity of multi-modal information has been fully utilized. To verify the effectiveness of MSSKF, experiments on the KITTI object detection benchmark dataset are conducted. It has been observed that the proposed method achieves more accurate detection for pedestrians and vehicles, with a 1.6% gain in AP(50) compared to the values of the original fusion method, reaching a score of 90.1%, and the mAP reached 60.9%. Experiments show that the proposed method introduces a new optimization idea for multi-modal fusion in the field of autonomous driving object detection, and the fusion detection efficiency is at over 12 fps on a single GPU.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
17  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
17  
ISSN:  
0263-2241  
eISSN:  
1873-412X  
入藏号:  
WOS:000806936500009  
Indexed Date:  
2022-06-26  
Institution Address:  
China Univ Min & Technol Beijing, Beijing 100083, Peoples R China  
Institution:

### 978. InDeandCoE: A framework based on multi-scale feature fusion and residual

摘要: Synthetic aperture radar (SAR) interferometry is a high-resolution microwave remote sensing imaging method. Over the past two decades, many researchers working on remote sensing have applied this technology in various disciplines, including environmental monitoring, disaster monitoring, and elevation mapping. However, due to the existence of many influencing factors in the acquisition stage, such as atmospheric humidity and temperature, the reflected wave signals from the ground will be disturbed when received by remote sensing satellites. The presence of noise in interferograms is inevitable. Therefore, the accuracy of interferometric SAR phase denoising and coherence estimation has a decisive impact on the validity of subsequent processing results. In this paper, we pioneer the use of a nested U-net as a feature extractor for interferometric SAR phase and coherence. In addition, we build a phase filter and a coherence estimator by using the residual learning module. With the aim of determining the unique non-local similarity of InSAR images, we use non-local convolution and channel attention mechanisms to extract features in different dimensions of the interferogram. Through quantitative and qualitative experiments, the proposed method performs better in phase denoising and coherence estimation than state-of-the-art methods.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0141-9382  
eISSN:  
1872-7387  
入藏号:  
WOS:001050328000001  
Indexed Date:  
2023-08-28  
Institution Address:  
Space Engn Univ, Sci & Technol Complex Elect Syst Simulat Lab, Beijing 101400, Peoples R China  
Institution Address:  
Beijing Inst Remote Sensing Informat, Beijing 100192, Peoples R China  
Institution Address:  
Beihang Univ, Beijing 100089, Peoples R China  
Institution Address:  
Naval Aviat Univ, Yantai 264001, Peoples R China  
Institution:  
Institution:  
Beijing Inst Remote Sensing Informat  
Institution:  
Institution:  
Naval Aviat Univ

### 979. Intelligent recognition algorithm and application of coal mine overhead

摘要: The intelligent recognition technology for Coal mine overhead passenger devices(Cmopd) plays a crucial role in achieving automated inspection, real-time monitoring, and warning tasks for cmopd, thereby promoting the intelligent development of coal mines. However, there are several challenges that need to be addressed, such as the limited number of samples in the cmopd dataset, poor lighting conditions in underground images, overlapping and occlusion between operating cmopd, varying sitting postures of miners, difficulty in detecting small cmopd targets, complex model deployment, and low efficiency of traditional recognition methods for cmopd with different passenger-carrying statuses.To overcome these challenges, a cmopd dataset was created from various coal mines in Guizhou province. The passenger-carrying status of cmopd was classified into two categories: cmopd with passengers (HC\_miner) and cmopd without passengers (HC\_nominer). The YOLOv8n single-stage object detection algorithm was used as the baseline model, and a coal mine cmopd intelligent recognition algorithm based on multi-scale feature fusion was proposed.In the image preprocessing stage, adaptive histogram equalization was employed to enhance image quality, and random rectangle masking was applied to simulate real scenarios where cmopd is occluded by underground objects during operation. This approach addressed the scarcity of cmopd image datasets and reduced the interference from negative underground environments. In the feature extraction stage, the partial convolution of the backbone network C2f module is replaced by deformable convolution, and a novel C2f\_DCN module is designed. This enhancement increased the dynamic adjustment capability of the target receptive field for cmopd with different passenger-carrying statuses, allowing the model to capture different scale information and better learn the coupled features of cmopd and miners. As a result, the model became more adaptable to various sitting postures of miners and improved its ability to identify cmopd targets with different passenger-carrying statuses. In the feature fusion stage, a path aggregation network with a coordinate attention mechanism (CLC-PAN-CA) was proposed to achieve cross-level contat of features and adaptively capture the contextual information of cmopd. The CLC-PAN-CA module effectively integrated multi-scale features and improved the accuracy of cmopd recognition. The experimental results show that the proposed model achieves a precision of 95.8%, which is 7.4% higher than the baseline model. The recall is 93.3%, representing an improvement of 9.8%, and the mean average precision is 95.6%, indicating a 7.7% increase. Furthermore, the model parameters and size are only 3.1\*10~6 and 6.1 MB, respectively. The recognition speed is 71 frames per second Compare to a variety of mainstream single-stage two-stage detection models, the proposed model demonstrated effective identification of cmopd targets with and without passengers, significantly improved the accuracy of cmopd recognition, reduced false positives and false negatives, and exhibited faster recognition speed and better extraction of contextual information. The proposed algorithm can meet the requirements of practical inspection scenarios and provide a feasible method for accurate recognition of cmopd with different passenger-carrying statuses. Finally, the proposed cmopd intelligent recognition algorithm and the underground monitoring video stream were embedded into the designed cmopd intelligent recognition system. Partial implementation approaches for deploying the video media stream into the cmopd intelligent recognition system were provided. The concept of an end-to-end integrated cmopd intelligent recognition system, which integrates the dispatching system on the ground and the monitoring system underground, was proposed. This increases the expectations for intelligent inspection applications in coal mines and provides real-time warnings for the safe transportation of cmopd with passengers.  
摘要:  
井下猴车智能识别技术是实现猴车自动化巡检、实时监测与预警任务的基础,是推动煤矿智能化发展的重要支撑。针对猴车数据集样本缺乏、井下图像光照环境较差、运行猴车之间存在重叠遮挡、矿工坐姿多变、小目标猴车难以检测、模型部署困难、不同载人状态井下猴车传统识别方法难度大效率低等问题,通过在贵州多个煤矿不同机段自建的猴车数据集,将猴车载人状态划分为载人猴车(HC\_miner)和无载人猴车(HC\_nominer)2种,提出了一种基于多尺度特征融合的井下猴车载人状态智能识别算法。图像预处理阶段,采用自适应直方图均衡以增强图像质量,通过随机矩形遮挡以模拟运行猴车被井下物体遮挡的真实场景,解决了猴车图像数据集体量匮乏的同时降低了井下负环境的干扰;特征提取阶段,将主干网络C2f模块部分卷积替换为可变形卷积(DCN),设计了一种C2f\_DCN模块,增加不同载人状态猴车目标感受野的动态调整能力以获取复杂多变的尺度信息,使模型更好地学习到猴车矿工的耦合特征及适应矿工各类坐姿细节,提升模型对不同载人状态猴车目标的辨识能力;特征融合阶段,提出了一种基于坐标注意力机制跨层级连接的路径聚合网络- CLC-PAN-CA模块,实现了深层网络与浅层网络特征间多尺度信息的复用,可自适应捕捉全局关键信息,建立网络之间的多尺度依赖,提升模型对小目标猴车重要特征的提取,减少背景噪声干扰,降低猴车目标漏检误检率。试验结果表明:提出模型的精确率为95.8%,对比基线模型提高了7.4%,召回率为93.3%,提高了9.8%,平均精度均值为95.6%,提升了7.7%,参数量和模型大小分别仅为3.1\*10~6和6.1 MB,识别速率为71帧。对比多种主流单阶段两阶段检测模型,提出模型可有效辨识有无载人猴车目标、显著提升井下猴车目标识别精度、减少漏检错检现象、具有较快的识别速度、更好的热度信息提取能力,可满足实际场景巡检需求,为不同载人状态的井下猴车精准识别提供了可行的方法。最后,将提出的猴车智能识别算法和井下监控视频流嵌入到设计的猴车智能识别系统中,构思了井上调度和井下监控 端到端一体化的猴车智能识别系统,增加了面向煤矿智能化巡检应用的期望值,可为井下猴车载人运输安全提供实时预警。  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0253-2336  
入藏号:  
CSCD:7880063  
Indexed Date:  
2025-03-15  
Institution Address:  
中国矿业大学(北京)应急管理与安全工程学院, 北京 100083, 中国  
Institution Address:  
贵州盘南煤炭开发有限责任公司, 盘州, 贵州 553000, 中国  
Institution Address:  
中国矿业大学(北京)应急管理与安全工程学院;;国家矿山安全监察局贵州局, ;;, ;;贵阳, 北京;;贵州 100083;;550001, 中国  
Institution Address:  
School of Emergency Management and Safety Engineering,China University of Mining and Technology-Beijing, Beijing 100083, China  
Institution Address:  
Guizhou Pannan Coal Development Co.,Ltd., Panzhou, 553000  
Institution Address:  
School of Emergency Management and Safety Engineering,China University of Mining and Technology-Beijing;;Guizhou Bureau of State Mine Safety Supervision Bureau, ;;, ;;Guiyang, Beijing;; 100083;;550001  
Institution:  
中国矿业大学(北京)应急管理与安全工程学院  
Institution:  
贵州盘南煤炭开发有限责任公司  
Institution:  
中国矿业大学(北京)应急管理与安全工程学院;;国家矿山安全监察局贵州局  
Institution:  
School of Emergency Management and Safety Engineering,China University of Mining and Technology-Beijing  
Institution:  
Guizhou Pannan Coal Development Co.,Ltd.  
Institution:  
School of Emergency Management and Safety Engineering,China University of Mining and Technology-Beijing;;Guizhou Bureau of State Mine Safety Supervision Bureau

### 980. Cross-modality synergy network for referring expression comprehension

摘要: Referring expression comprehension and segmentation aim to locate and segment a referred instance in an image according to a natural language expression. However, existing methods tend to ignore the interaction between visual and language modalities for visual feature learning, and establishing a synergy between the visual and language modalities remains a considerable challenge. To tackle the above problems, we propose a novel end-to-end framework, Cross-Modality Synergy Network (CMS-Net), to address the two tasks jointly. In this work, we propose an attention-aware representation learning module to learn modal representations for both images and expressions. A language self-attention submodule is proposed in this module to learn expression representations by leveraging the intra-modality relations, and a language-guided channel-spatial attention submodule is introduced to obtain the language aware visual representations under language guidance, which helps the model pay more attention to the referent-relevant regions in the images and relieve background interference. Then, we design a cross-modality synergy module to establish the inter-modality relations for modality fusion. Specifically, a language-visual similarity is obtained at each position of the visual feature map, and the synergy is achieved between the two modalities in both semantic and spatial dimensions. Furthermore, we propose a multi-scale feature fusion module with a selective strategy to aggregate the important information from multi-scale features, yielding target results. We conduct extensive experiments on four challenging benchmarks, and our framework achieves significant performance gains over state-of-the-art methods. (c) 2021 Elsevier B.V. All rights reserved.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
16  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
16  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000710121100009  
Indexed Date:  
2021-11-03  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Management & Control Complex Syst, Beijing, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Artificial Intelligences, Beijing, Peoples R China  
Institution:  
Institution:

### 981. Two-Stage Segmentation Algorithm of High Resolution Skin Melanoma Image

摘要: Skin melanoma imaging faces a few challengesthe required resolution is often large,the pathological features of skin melanoma are diverse,and many segmentation algorithms are not sufficiently accurate and require huge amounts of GPU memory. Aiming at this problem,this paper proposes a low GPU memory consumption algorithm based on two-stage fine segmentation. The global segmentation network adopted in the first stage of the algorithm improves the feature pyramid structure with ResNet50 as the backbone. During image feature extraction,the Global Pyramid Average Pooling Module(GPAPM)is used to enhance the extraction of image global semantic information. The multi-scale feature fusion branch is adopted to integrate the semantic information of the high-level feature map into the lower-level feature map to enhance the representation ability of the semantic information of the low-level feature map. In the second stage,a global to local fine segmentation strategy is adopted. The image is clipped based on the global segmentation results to reduce the candidate area,and the image is then input into the local segmentation network. The local segmentation network only processes pixels in the candidate region and shares image features with the corresponding layer of the global network. The GPU memory consumption is minimized while the segmentation results are improved. The algorithm is verified on the classic dataset ISIC2018. The experimental results show that the accuracy and IOU of the algorithm reach 93.5% and 82.1%,respectively,and compared with other classical segmentation algorithms,the algorithm has the highest accuracy and reduces the GPU memory by 22.8% to 36.9%,indicating its effective applicability to the segmentation of high resolution skin melanoma images.  
摘要:  
皮肤黑色素瘤切片图像分辨率过大且病理特征表现形式多样,现有很多分割算法结果不精准同时消耗巨大显卡内存。针对该问题,提出一种低显存消耗的两阶段式精细分割算法。该算法第一阶段采用全局分割网络对以ResNet50为骨干的特征金字塔结构进行改进,图像特征提取过程中使用全局金字塔平均池化模块增强图像全局语义信息的提取,并采用多尺度特征融合分支将高层特征图的语义信息融入到低层特征图中,增强低层特征图语义信息的表征能力。第二阶段采用一种全局到局部的精细分割策略,以全局分割结果为基准对图像进行剪裁,得到一个较小的候选区域,将其输入到局部分割网络中,局部分割网络仅处理候选区域内的像素并与全局网络对应层共享图像特征,精细分割结果的同时减少显存的消耗。在经典数据集ISIC2018上的实验结果显示,该算法的准确度和IOU分别达到93.5%和82.1%,相较于对比的经典分割算法精度最高且占用的显卡内存减少了22.8%~ 36.9%,能有效适用于高分辨率皮肤病灶图像的分割任务。  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1000-3428  
入藏号:  
CSCD:7593411  
Indexed Date:  
2024-01-19  
Institution Address:  
兰州理工大学计算机与通信学院, 兰州, 甘肃 730050, 中国  
Institution Address:  
School of Computer and Communication,Lanzhou University of Technology, Lanzhou, Gansu 730050, China  
Institution:  
兰州理工大学计算机与通信学院  
Institution:  
School of Computer and Communication,Lanzhou University of Technology

### 982. MSEF-Net: Multi-scale edge fusion network for lumbosacral plexus

摘要: Nerve damage of spine areas is a common cause of disability and paralysis. The lumbosacral plexus segmentation from magnetic resonance imaging (MRI) scans plays an important role in many computer -aided diagnoses and surgery of spinal nerve lesions. Due to the complex structure and low contrast of the lumbosacral plexus, it is difficult to delineate the regions of edges accurately. To address this issue, we propose a Multi -Scale Edge Fusion Network (MSEF-Net) to fully enhance the edge feature in the encoder and adaptively fuse multiscale features in the decoder. Specifically, to highlight the edge structure feature, we propose an edge feature fusion module (EFFM) by combining the Sobel operator edge detection and the edge -guided attention module (EAM), respectively. To adaptively fuse the multi -scale feature map in the decoder, we introduce an adaptive multi -scale fusion module (AMSF). Our proposed MSEF-Net method was evaluated on the collected spinal MRI dataset with 89 patients (a total of 2848 MR images). Experimental results demonstrate that our MSEF-Net is effective for lumbosacral plexus segmentation with MR images, when compared with several state-of-the-art segmentation methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0933-3657  
eISSN:  
1873-2860  
入藏号:  
WOS:001167882700001  
PubMed ID:  
38325928  
Indexed Date:  
2024-03-06  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Comp Sci & Technol, Minist Educ, Key Lab Brain Machine Intelligence Technol, Nanjing 211106, Peoples R China  
Institution Address:  
Shandong Univ, Qilu Hosp, Cheeloo Coll Med, Dept Orthoped, Jinan 250012, Peoples R China  
Institution Address:  
Shandong Prov Hosp, Dept Med Imaging, Jinan 250021, Peoples R China  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Shenzhen Res Inst, Shenzhen 518063, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 983. MSA-Net: Multiscale spatial attention network for medical image

摘要: Background: Edge accuracy and positional accuracy are the two goals pursued by med-ical image segmentation. In clinical medicine diagnosis and research, these two goals enable medical image segmentation techniques to help in the effective determination of lesions and lesion analysis. At present, U-Net has become the most important network in the field of image segmentation, and the technologies used in various achievements are derived from its architecture, which also proves from practice that the network structure proposed by U-Net is effective. Objective: We have found in a large number of experiments that classical networks indeed show good performance in the field of medical segmentation, but there are still some deficiencies in edge determination and network robustness, especially in the face of blurred edges, the processing results often fail to achieve the expected results. In order to be able to locate segmentation targets and achieve effective determination of blurred edges, a Multiscale Spatial Attention Network (MSA-Net) is proposed as in Fig. 1. Method: In MSA-Net, the Multiscale Pyramid Attention Block (MPAB) is created to enhance the capture of high-level semantic information. In addition, the network uses ASPP, which not only expands the network's field of view, but also captures richer feature information. In the decoding phase, the Feature Fusion Block (FFB) is created to enable better focus on different dimensional information features and to enhance the feature fusion process. Result: To demonstrate the effectiveness of the network, we validate the performance of MSA-Net on four datasets (ISIC2016, DSB2018, JSRT, GlaS) in three different categories. Compared with mainstream networks, MSA-Net shows better results in detail features, target localization, and edge processing. Finally, we also demonstrate the effectiveness of the MSA-Net architecture through ablation experiments. (c) 2023 THE AUTHORS. Published by Elsevier BV on behalf of Faculty of Engineering, Alexandria University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/ licenses/by-nc-nd/4.0/).  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
12  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
12  
ISSN:  
1110-0168  
eISSN:  
2090-2670  
入藏号:  
WOS:000962660100001  
Indexed Date:  
2023-04-28  
Institution Address:  
Shandong Technol & Business Univ, Sch Informat & Elect Engn, Yantai 264005, Peoples R China  
Institution Address:  
Shandong Technol & Business Univ, Sch Comp Sci & Technol, Yantai 264005, Peoples R China  
Institution Address:  
Coinnovat Ctr Shandong Coll & Univ Future Intellig, Yantai 264005, Peoples R China  
Institution:  
Institution:  
Institution:  
Coinnovat Ctr Shandong Coll & Univ Future Intellig

### 984. A Deep Segmentation Network of Multi-Scale Feature Fusion Based on

摘要: Recently, coronary heart disease has attracted more and more attention, where segmentation and analysis for vascular lumen contour are helpful for treatment. And intravascular optical coherence tomography (IVOCT) images are used to display lumen shapes in clinic. Thus, an automatic segmentation method for IVOCT lumen contour is necessary to reduce the doctors' workload while ensuring diagnostic accuracy. In this paper, we proposed a deep residual segmentation network of multi-scale feature fusion based on attention mechanism (RSM-Network, Residual Squeezed Multi-Scale Network) to segment the lumen contour in IVOCT images. Firstly, three different data augmentation methods including mirror level turnover, rotation and vertical flip are considered to expand the training set. Then in the proposed RSM-Network, U-Net is contained as the main body, considering its characteristic of accepting input images with any sizes. Meanwhile, the combination of residual network and attention mechanism is applied to improve the ability of global feature extraction and solve the vanishing gradient problem. Moreover, the pyramid feature extraction structure is introduced to enhance the learning ability for multi-scale features. Finally, in order to increase the matching degree between the actual output and expected output, the cross entropy loss function is also used. A series of metrics are presented to evaluate the performance of our proposed network and the experimental results demonstrate that the proposed RSM-Network can learn the contour details better, contributing to strong robustness and accuracy for IVOCT lumen contour segmentation.  
Web of Science 核心合集中的 "被引频次":  
33  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
5  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
35  
ISSN:  
1545-5963  
eISSN:  
1557-9964  
入藏号:  
WOS:000615042600007  
PubMed ID:  
32078556  
Indexed Date:  
2021-12-14  
Institution Address:  
Xiamen Univ, Sch Informat, Xiamen 361005, Fujian, Peoples R China  
Institution Address:  
Tongji Univ, Dept Comp Sci & Technol, Shanghai 201804, SH, Peoples R China  
Institution Address:  
Tongji Univ, Coll Elect & Informat Engn, Shanghai 201804, SH, Peoples R China  
Institution Address:  
Univ Essex, Sch Comp Sci & Elect Engn, Colchester CO4 3SQ, Essex, England  
Institution Address:  
Xiamen Univ, Sch Informat, Dept Artificial Intelligence, Media Analyt & Comp Lab, Xiamen 361005, Fujian, Peoples R China  
Institution Address:  
Xiamen Univ, Sch Aerosp Engn, Xiamen 361005, Fujian, Peoples R China  
Institution Address:  
Nanyang Technol Univ, Sch Mech & Aerosp Engn, Singapore 639798, Singapore  
Institution Address:  
Univ Sunderland, Fac Comp Sci, Sunderland SR6 0DD, England  
Institution Address:  
Tongji Univ, Shanghai Tongji Hosp, Dept Cardiol, Shanghai 200065, SH, Peoples R China  
Institution Address:  
Tongji Univ, Sch Software Engn, Shanghai 201804, SH, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 985. X-Ray Image Controlled Knife Detection and Recognition Based on Improved

摘要: In the automatic X-ray imaging systems used to identify knives in security check, using the original single shot multibox detector (SSD) algorithm, the shallow feature maps are poorly represented, features of small targets gradually disappear during the training stage, leading to low detection accuracy and poor real-time performance, and the small targets such as the controlled knives in security check are missing and checked out by mistake. To solve this problem, the original SSD was improved in two ways. On the one hand, the SSD-Resnet34 network model was constructed by replacing the basic network VGG16 in the SSD using a ResNet34 network with stronger anti degradation performance, and the last three layers of the basic network were convolved and a new low-level feature map was created by lightweight network fusion. Part of the extended layer of the network was deconvolved to form a new high-level feature map. On the other hand, jumping connection was adopted to achieve multi-scale feature fusion between the high-level feature map and the low-level feature map. Analysis of test data shows that the improved algorithm demonstrates improved detection speed and detection accuracy of small targets, such as the Xray image controlled knives. And the algorithm demonstrates improved robustness and high real-time performance. Using the VOC2007+2012 general dataset, the detection accuracy of the improved SSD algorithm is 1.7% higher than that of the SSD algorithm, reaching 80.5%  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
7  
ISSN:  
1006-4125  
入藏号:  
WOS:000686478600006  
Indexed Date:  
2021-08-29  
Institution Address:  
Xian Polytech Univ, Coll Elect & Informat, Xian 710048, Shaanxi, Peoples R China  
Institution:

### 986. A medical assistant segmentation method for MRI images of osteosarcoma

摘要: Nowadays, the most common primary bone tumor is osteosarcoma, which mostly occurs in teenagers. A common diagnosis method is currently that doctors manually diagnose osteosarcoma in magnetic resonance imaging (MRI) images because it is nonradioactive and has no biological damage to brain tissue and more obvious performance in soft tissue components such as tumors, blood vessels, and muscles in MRI images. However, this method is labor-intensive and time-consuming work, and cannot guarantee the accuracy of the diagnostic results. Existing osteosarcoma MRI image segmentation methods either aim to model the global context to improve the inner consistency of objects, or multiscale feature fusion to refine the detail of objects along their boundaries, which all ignore the interaction between the body of the object and the object boundary. Therefore, this paper proposes a novel segmentation method for osteosarcoma MRI images based on DecoupleSegNet, which explores the relationship between body feature and edge feature. It can assist doctors in diagnosing osteosarcoma and improve their work efficiency. First, we warp the feature of MRI images through learning a flow field so we can make the object more consistent. We then make further work to optimize the resulting body feature and residual edge feature through explicitly sampling pixels from different parts under decoupled supervision. Through these steps, we finally obtain the final feature map with fine boundaries from the MRI image of osteosarcoma. We take a test by using more than 80,000 osteosarcoma MRI images obtained from three hospitals in China. We find that compared with existing osteosarcoma MRI image segmentation methods, our proposed method achieves 90.51 Intersection of Union % with few parameters on the test, outperforming other models. In the test, we prove that our proposed method has better accuracy and lower resource consumption.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
41  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
41  
ISSN:  
0884-8173  
eISSN:  
1098-111X  
入藏号:  
WOS:000820954100001  
Indexed Date:  
2022-07-10  
Institution Address:  
Cent South Univ, Sch Comp Sci & Engn, Changsha, Peoples R China  
Institution Address:  
Monash Univ, Res Ctr Artificial Intelligence, Clayton, Vic, Australia  
Institution Address:  
Cent South Univ, Xiangya Hosp 2, Dept Spine Surg, Changsha 410011, Peoples R China  
Institution:  
Institution:  
Institution:

### 987. rU-Net, Multi-Scale Feature Fusion and Transfer Learning: Unlocking the

摘要: This study introduces an innovative deep-learning model for cuffless blood pressure estimation using PPG and ECG signals, demonstrating state-of-the-art performance on the largest clean dataset, PulseDB. The rU-Net architecture, a fusion of U-Net and ResNet, enhances both generalization and feature extraction accuracy. Accurate multi-scale feature capture is facilitated by short-time Fourier transform (STFT) time-frequency distributions and multi-head attention mechanisms, allowing data-driven feature selection. The inclusion of demographic parameters as supervisory information further elevates performance. On the calibration-based dataset, our model excels, achieving outstanding accuracy (SBP MAE +/- std: 4.49 +/- 4.86 mmHg, DBP MAE +/- std: 2.69 +/- 3.10 mmHg), surpassing AAMI standards and earning a BHS Grade A rating. Addressing the challenge of calibration-free data, we propose a fine-tuning-based transfer learning approach. Remarkably, with only 10% data transfer, our model attains exceptional accuracy (SBP MAE +/- std: 4.14 +/- 5.01 mmHg, DBP MAE +/- std: 2.48 +/- 2.93 mmHg). This study sets the stage for the development of highly accurate and reliable wearable cuffless blood pressure monitoring devices.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2168-2194  
eISSN:  
2168-2208  
入藏号:  
WOS:001394517600007  
PubMed ID:  
39423074  
Indexed Date:  
2025-01-15  
Institution Address:  
Fudan Univ, Sch Informat Sci & Technol, Shanghai 200437, Peoples R China  
Institution Address:  
Greater Bay Area Inst Precis Med, Guangzhou 511462, Peoples R China  
Institution Address:  
Guangdong Univ Technol, Sch Informat Engn, Guangzhou 510006, Peoples R China  
Institution Address:  
Fudan Univ, Huashan Hosp, Sch Life Sci, Dept Dermatol, Shanghai 200437, Peoples R China  
Institution:  
Institution:  
Greater Bay Area Inst Precis Med  
Institution:  
Institution:

### 988. End-to-end residual attention mechanism for cataractous retinal image

摘要: Background and objective: : Cataract is one of the most common causes of vision loss. Light scattering due to clouding of the lens in cataract patients makes it extremely difficult to image the retina of cataract patients with fundus cameras, resulting in a serious decrease in the quality of the retinal images taken. Furthermore, the age of cataract patients is generally too old, in addition to cataracts, the patients often have other retinal diseases, which brings great challenges to experts in the clinical diagnosis of cataract patients using retinal imaging.  
Methods: : In this paper, we present the End-to-End Residual Attention Mechanism (ERAN) for Cataractous Retinal Image Dehazing, which it includes four modules: encoding module, multi-scale feature extraction module, feature fusion module, and decoding module. The encoding module encodes the input cataract haze image into an image, facilitating subsequent feature extraction and reducing memory usage. The multi-scale feature extraction module includes a hole convolution module, a residual block, and an adaptive skip connection, which can expand the receptive field and extract features of different scales through weighted screening for fusion. The feature fusion module uses adaptive skip connections to enhance the network's ability to extract haze density images to make haze removal more thorough. Furthermore, the decoding module performs non-linear mapping on the fused features to obtain the haze density image, and then restores the haze-free image.  
Results: : The experimental results show that the proposed method has achieved better objective and subjective evaluation results, and has a better dehazing effect.  
Conclusion: : We proposed ERAN method not only provides visually better images, but also helps experts better diagnose other retinal diseases in cataract patients, leading to better care and treatment. (C) 2022 Elsevier B.V. All rights reserved.  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
0169-2607  
eISSN:  
1872-7565  
入藏号:  
WOS:000821201900010  
PubMed ID:  
35397410  
Indexed Date:  
2022-07-18  
Institution Address:  
China Univ Min & Technol, Engn Res Ctr Intelligent Control Underground Spac, Minist Educ, Xuzhou 221116, Jiangsu, Peoples R China  
Institution Address:  
China Univ Min & Technol, Sch Informat & Control Engn, Xuzhou 221116, Jiangsu, Peoples R China  
Institution:  
Institution:

### 989. RT-DETRmg: a lightweight real-time detection model for small traffic

摘要: In intelligent transportation systems, real-time detection performance and accuracy are essential metrics. This paper proposes a lightweight real-time detection model, RT-DETRmg, to address the challenges of false and missed detections of small traffic signs and to improve the algorithm's real-time performance. RT-DETRmg enhances the multi-scale feature extraction capability of the RT-DETR backbone network by incorporating a Multiple Scale Sequence Fusion module, which effectively integrates global and local semantic information from different scales of images. Additionally, a cascaded group attention module is utilized within an efficient hybrid encoder to reduce computational complexity, thereby enhancing real-time performance. To further optimize small object detection, a small receptive field feature layer is introduced, while a large receptive field feature layer is removed. Experimental results on the TT100K and GTSDB datasets demonstrate the superiority of RT-DETRmg over existing models. On the TT100K dataset, RT-DETRmg achieves a 2.0% improvement in mean average precision and a 6.6% increase in frames per second compared to the baseline RT-DETR model, while reducing model parameters and computational complexity. On the GTSDB dataset, RT-DETRmg further demonstrates its strong generalization ability, achieving a 2.2% improvement in the F1 score and a 1.7% increase in mean average precision compared to the baseline network. These findings highlight the effectiveness of RT-DETRmg in enhancing both detection accuracy and real-time performance of small traffic signs in diverse scenarios.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0920-8542  
eISSN:  
1573-0484  
入藏号:  
WOS:001380058600001  
Indexed Date:  
2024-12-24  
Institution Address:  
Southwest Petr Univ, Chengdu 610500, Sichuan, Peoples R China  
Institution Address:  
State Grid Sichuan Informat & Telecommun Co, Chengdu 610095, Sichuan, Peoples R China  
Institution:  
Institution:  
State Grid Sichuan Informat & Telecommun Co

### 990. MSLAENet: Multiscale Learning and Attention Enhancement Network for

摘要: The effective use of multimodal data to obtain accurate land cover information has become an interesting and challenging research topic in the field of remote sensing. In this article, we propose a new method, multiscale learning and attention enhancement network (MSLAENet), to implement hyperspectral image (HSI) and light detection and ranging (LiDAR) data fusion classification in an end-to-end manner. Specifically, our model consists of three main modules. First, we design the composite attention module, which adopts self-attention to enhance the feature representations of HSI and LiDAR data, respectively, and cross-attention to achieve cross-modal information enhancement. Second, the proposed multiscale learning module combines self-calibrated convolutions and hierarchical residual structure to extract different scales of information to further improve the representation capability of the model. Finally, the attention-based feature fusion module fully considers the complementary information properties between different modalities and adaptively fuses heterogeneous features from different modalities. To test the performance of MSLAENet, we conduct experiments on three multimodal remote sensing datasets and compare them with the state-of-the-art fusion model, which demonstrates the effectiveness and superiority of the model.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
15  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
15  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000914641600005  
Indexed Date:  
2023-02-07  
Institution Address:  
Xinjiang Univ, Coll Informat Sci & Engn, Urumqi 830008, Peoples R China  
Institution Address:  
Key Lab Signal Detect & Proc Xinjiang Uygur Autono, Urumqi 830014, Peoples R China  
Institution Address:  
Key Lab Software Engn, Urumqi 830008, Peoples R China  
Institution Address:  
Key Lab Software Engn, Coll Software, Urumqi 830008, Peoples R China  
Institution:  
Institution:  
Key Lab Signal Detect & Proc Xinjiang Uygur Autono  
Institution:  
Key Lab Software Engn  
Institution:  
Key Lab Software Engn

### 991. Multiple lesion segmentation in diabetic retinopathy with dual-input

摘要: To address the issue of complex structure, various sizes and the interclass similarity of different lesions, this paper proposes a dual-input attentive RefineNet (DARNet) for automatic multiple lesion segmentation of diabetic retinopathy. DARNet includes a global image encoder, local image encoder and attention refinement decoder. The whole image and the patch image are used as the dual input and fed into ResNet50 and ResNet101 for down-sampling, respectively. The high-level attention refinement decoder adopts a dual attention mechanism to integrate the same-level features in the two encoders with the output of the low-level attention refinement module for multiscale feature fusion, which focuses the model on the lesion area to generate accurate predictions. We evaluated the segmentation performance of four lesions on three datasets, and the proposed method reached an average accuracy of 0.9582/0.9617/0.9578 and a dice score of 0.9521/0.9637/0.9508 on IDRiD, E-ophtha and DDR. Extensive experimental results demonstrate the proposed DARNet outperforms the state- of-the-art models and has better robustness and accuracy. It not only preserves the contour details and shape features of multiscale lesions, but also overcomes the interference of similar tissues and noises to realize accurate multi-lesion segmentation.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000765733600001  
Indexed Date:  
2022-03-18  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao 266590, Shandong, Peoples R China  
Institution Address:  
Shandong Univ Sci & Technol, Shandong Prov Key Lab Wisdom Min Informat Technol, Qingdao 266590, Shandong, Peoples R China  
Institution:  
Institution:

### 992. Enhancing Landslide Detection: A Novel LA-YOLO Model for

摘要: As a geological disaster widely distributed in the southern regions of China, rainfall-induced shallow landslides pose a significant threat to affected areas. Timely detection of landslides is crucial in the effective response to such disasters. However, landslide detection faces adverse impacts from various factors, such as insufficient sample data, complex model structures, and limitations in detection accuracy during the actual detection process. In this study, high-quality image samples were collected from multiple landslide disaster areas in southern China, and a rainfall-induced shallow landslide sample database was constructed in the region. Based on this, a lightweight attention-guided YOLO model (LA-YOLO) was proposed to improve the detection performance of YOLO model for rainfall-induced shallow landslides. First, CG block is introduced to enhance the C2f module, enriching the feature representation capability through multiscale feature fusion and reducing the model's parameters and computational complexity. Second, the SimAM attention module is used to focus on the target regions, improving feature extraction effectiveness. Experimental results show that the model parameters of LA-YOLO were reduced by approximately 30%, with precision, recall, and mean average precision (mAP) on the landslide sample dataset increasing by 2.6%, 0.7%, and 2.2%, respectively. While ensuring model detection performance, the model structure was significantly optimized, achieving both lightweight and accuracy goals, confirming the model's superiority in monitoring rainfall-induced shallow landslide disasters.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001445059100009  
Indexed Date:  
2025-03-21  
Institution Address:  
Fuzhou Univ, Coll Environm & Safety Engn, Fuzhou 350108, Peoples R China  
Institution Address:  
Fuzhou Univ, Acad Digital China Fujian, Fuzhou 350108, Peoples R China  
Institution Address:  
Fuzhou Univ, Dept Geotech & Geol Engn, Fuzhou 350108, Peoples R China  
Institution:  
Institution:  
Institution:

### 993. Memory Based Temporal Fusion Network for Video Deblurring

摘要: Video deblurring is one of the most challenging vision tasks because of the complex spatial-temporal relationship and a number of uncertainty factors involved in video acquisition. As different moving objects in the video exhibit different motion trajectories, it is difficult to accurately capture their spatial-temporal relationships. In this paper, we proposed a memory-based temporal fusion network (TFN) to capture local spatial-temporal relationships across the input sequence for video deblurring. Our temporal fusion network consists of a memory network and a temporal fusion block. The memory network stores the extracted spatial-temporal relationships and guides the temporal fusion blocks to extract local spatial-temporal relationships more accurately. In addition, in order to enable our model to more effectively fuse the multiscale features of the previous frame, we propose a multiscale and multi-hop reconstruction memory network (RMN) based on the attention mechanism and memory network. We constructed a feature extractor that integrates residual dense blocks with three downsample layers to extract hierarchical spatial features. Finally, we feed these aggregated local features into a reconstruction module to restore sharp video frames. Experimental results on public datasets show that our temporal fusion network has achieved a significant performance improvement in terms of PSNR metrics (over 1dB) over existing state-of-the-art video deblurring methods.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0920-5691  
eISSN:  
1573-1405  
入藏号:  
WOS:000971433900001  
Indexed Date:  
2023-05-06  
Institution Address:  
Xidian Univ, Xian, Peoples R China  
Institution Address:  
West Virginia Univ, Morgantown, WA USA  
Institution:  
Institution:  
West Virginia Univ

### 994. DFSA-DAN: dynamic fusion of statistical metric and adversarial learning

摘要: The advancement of deep transfer learning has motivated research into the realization of intelligent fault diagnosis schemes for rolling bearing. Nevertheless, existing research rarely provides further insight into the importance of statistical distance metric-based methods and adversarial learning-based methods in domain adaptation, and the commonly used feature extractors are more difficult to extract features suitable for domain transformation. In this paper, a dynamic fusion of statistical metric and adversarial learning for domain adaptation network is proposed to achieve a dynamic measure of the importance of different domain adaptation methods. This new model utilizes a local maximum mean discrepancy metric to adjust the conditional distribution and adversarial training to adjust the marginal distribution between domains. Meanwhile, to assess the importance of the two distributions, a dynamic adaptation factor is introduced for dynamic evaluation. In addition, to extract features that are more suitable for domain transformation, the model incorporates a dual depth convolutional path with an attention mechanism as a feature extractor, enabling multi-scale feature extraction. Experimental results demonstrate the model's superior generalization capability and robustness, enabling effective cross-domain fault diagnosis in diverse scenarios.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001220796800001  
Indexed Date:  
2024-05-18  
Institution Address:  
Hangzhou Dianzi Univ, Sch Elect & Informat, Hangzhou 310018, Peoples R China  
Institution:

### 995. BIBED-Seg: Block-in-Block Edge Detection Network for Guiding Semantic

摘要: Edge optimization of semantic segmentation results is a challenging issue in remote sensing image processing. This article proposes a semantic segmentation model guided by a block-in-block edge detection network named BIBED-Seg. This is a two-stage semantic segmentation model, where edges are extracted first and then segmented. We do two key works: The first work is edge detection, and we present BIBED, a block-in-block edge detection network, to extract the accurate boundary features. Here, the edge detection of multiscale feature fusion is first realized by creating the block-in-block residual network structure and devising the multilevel loss function. Second, we add the channel and spatial attention module into the residual structure to improve high-resolution remote sensing images' boundary positioning and detection accuracy by focusing on their channel and spatial dimensions. Finally, we evaluate our method on International Society for Photogrammetry and Remote Sensing (ISPRS) Potsdam and Vaihingen data sets and obtain ODS F-measure of 0.6671 and 0.7432, higher than other excellent edge detection methods. The second work is two-stage segmentation. First, the proposed BIBED is individually pretrained, and subsequently, the pretrained model is introduced into the entire segmentation network to extract boundary features. In the second segmentation stage, the edge detection network is used to constrain semantic segmentation results by loss cycles and feature bootstrapping. Our best model obtains the OA of 90.2%, 87.7%, and 81.5%, the IOU of 76.0%, 69.6%, and 61.3% on the ISPRS and WHDLD datasets, respectively.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
19  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
19  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000935002300001  
Indexed Date:  
2023-03-18  
Institution Address:  
Southwest Jiaotong Univ, Fac Geosci & Environm Engn, Chengdu 611756, Peoples R China  
Institution:

### 996. NAS-Kernel: Learning Suitable Gaussian Kernel for Remote-Sensing Object

摘要: The purpose of object counting is to estimate the number of specific kinds of objects in a given image. In remote-sensing imagery, challenges arise in object counting due to issues like scale variations and complex backgrounds. Existing density map-based object counting methods have achieved satisfactory performance in some general scenarios (i.e., crowd counting and vehicle counting) and have become the mainstream methods. These density map-based counting methods use a fixed Gaussian kernel in the density map generation stage, thus they are not well adapted to the challenges such as scale variations present in remote-sensing scenes. In this letter, we propose to use the strategy of neural architecture search (NAS-Kernel) to select appropriate Gaussian kernels corresponding to objects of different scales in the Gaussian density map generation stage. NAS-Kernel is a plug-and-play algorithm that can be used in other density map-based counting methods. In addition, a contextual path aggregation (CPA) feature fusion strategy is proposed to fuse multiscale feature information. The ablation experiments verify that the proposed method can significantly improve the performance of the baseline. Experimental results on the four subdatasets of RSOC show that the proposed method achieves state-of-the-art performance. On the Building subdataset, the proposed method achieves 18% and 12% lower mean absolute error (MAE) and root mean square error (RMSE) than the existing methods.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001079253100005  
Indexed Date:  
2023-10-22  
Institution Address:  
Northwestern Polytech Univ, Sch Artificial Intelligence OPt & Elect iOPEN, Xian 710072, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Sch Comp Sci, Xian 710072, Peoples R China  
Institution Address:  
Minist Ind & Informat Technol, Key Lab Intelligent Interact & Applicat, Xian 710072, Peoples R China  
Institution Address:  
Shanghai Artificial Intelligence Lab, Shanghai 200232, Peoples R China  
Institution:  
Institution:  
Institution:  
Minist Ind & Informat Technol  
Institution:  
Shanghai Artificial Intelligence Lab

### 997. ImLiDAR: Cross-Sensor Dynamic Message Propagation Network for 3-D Object

摘要: LiDAR and camera, as two different sensors, supply geometric (point clouds) and semantic (RGB images) information of 3-D scenes. However, it is still challenging for existing methods to fuse data from the two cross sensors, making them complementary for quality 3-D object detection (3OD). We propose ImLiDAR, a new 3OD paradigm to narrow the cross-sensor discrepancies by progressively fusing the multiscale features of camera Images and LiDAR point clouds. ImLiDAR enables to provide the detection head with cross-sensor yet robustly fused features. To achieve this, two core designs exist in ImLiDAR. First, we propose a cross-sensor dynamic message propagation (CDMP) module to combine the best of the multiscale image and point features. Second, we raise a direct set prediction problem that allows designing an effective set-based detector (SD) to tackle the inconsistency of the classification and localization confidences, and the sensitivity of hand-tuned hyperparameters. Besides, the novel SD can be detachable and easily integrated into various detection networks. Comparisons on the KITTI, nuScenes, and SUN-RGBD datasets all show clear visual and numerical improvements of our ImLiDAR over 45 state-of-the-art 3OD methods.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001087759000009  
Indexed Date:  
2023-11-12  
Institution Address:  
Wuhan Univ, Sch Cyber Sci & Engn, Minist Educ, Key Lab Aerosp Informat Secur & Trusted Comp, Wuhan 430072, Peoples R China  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Shenzhen Res Inst, Shenzhen 211106, Peoples R China  
Institution Address:  
Lingnan Univ, Dept Comp & Decis Sci, Hong Kong, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Sch Nursing, Hong Kong, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 998. 被撤回的出版物: Multi-scale fusion public gathering recognition based on

摘要: With the rapid economic development and high concentration of urban population, people's income level and quality of life continue to improve, resulting in more and more crowded scenes caused by people going out. Especially in urban commercial centers, transportation hubs, sports venues during important events, tourist attractions, etc., crowd gatherings occur frequently. However, accidents involving crowd gatherings in public places occur frequently, causing heavy casualties and property losses. Therefore, for crowd recognition, this paper proposes a new method to accurately estimate the number of dense crowds. In this method, a density map with accurate pedestrian locations is first generated using the focal inverse distance transform and used as ground truth labels for network training. Then, a multi-scale feature fusion algorithm based on residual network is designed, combining spatial and channel attention mechanisms to improve the accuracy and stability of crowd density estimation. In dense crowds, the phenomenon of overlapping and occlusion of people is very common and serious, making it difficult for existing pedestrian detection methods to distinguish each individual and accurately count the flow of people. To solve this problem, this paper proposes a density map-based method that uses a local maximum detection strategy and a K-nearest neighbor algorithm to convert the density map into the corresponding dense head bounding box. This method can effectively reduce the impact of occlusion and improve the accuracy of people counting. In order to further improve the estimation accuracy, a pattern recognition density peak clustering algorithm is introduced to study the clustered crowds. By treating the head bounding box as an element point, the distance between each element point is calculated, and the density of each point is calculated. Then perform clustering to find the cluster center with the highest density in each class. Finally, by comparing the density of each cluster center with the corresponding density threshold and adopting the corresponding decision-making method, the accuracy of people counting is further improved.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1064-1246  
eISSN:  
1875-8967  
入藏号:  
WOS:001193319500050  
Indexed Date:  
2024-04-30  
Institution Address:  
Anhui Univ, Sch Artificial Intelligence, Hefei, Peoples R China  
Institution Address:  
Lyceum Philippines Univ, Grad Sch, Manila 1002, Philippines  
Institution Address:  
Jinling High Sch, Xishan Campus, Nanjing, Peoples R China  
Institution:  
Institution:  
Institution:  
Jinling High Sch

### 999. On the regularization of feature fusion and mapping for fast MR

摘要: Multi-contrast (MC) Magnetic Resonance Imaging (MRI) of the same patient usually requires long scanning times, despite the images sharing redundant information. In this work, we propose a new iterative network that utilizes the sharable information among MC images for MRI acceleration. The proposed network has reinforced data fidelity control and anatomy guidance through an iterative optimization procedure of Gradient Descent, leading to reduced uncertainties and improved reconstruction results. Through a convolutional network, the new method incorporates a learnable regularization unit that is capable of extracting, fusing, and mapping shareable information among different contrasts. Specifically, a dilated inception block is proposed to promote multi-scale feature extractions and increase the receptive field diversity for contextual information incorporation. Lastly, an optimal MC information feeding protocol is built through the design of a complementary feature extractor block. Comprehensive experiments demonstrated the superiority of the proposed network, both qualitatively and quantitatively.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
18  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
被引频次合计:  
19  
ISSN:  
0730-725X  
eISSN:  
1873-5894  
入藏号:  
WOS:000617062100006  
PubMed ID:  
33400936  
Indexed Date:  
2021-01-12  
Institution Address:  
Univ Queensland, Sch Informat Technol & Elect Engn, Brisbane, Qld, Australia  
Institution Address:  
Griffith Univ, Sch Informat & Commun Technol, Brisbane, Qld, Australia  
Institution:  
Institution:

### 1000. LPE-Unet: An Improved UNet Network Based on Perceptual Enhancement

摘要: In Computed Tomography (CT) images of the coronary arteries, the segmentation of calcified plaques is extremely important for the examination, diagnosis, and treatment of coronary heart disease. However, one characteristic of the lesion is that it has a small size, which brings two difficulties. One is the class imbalance when computing loss function and the other is that small-scale targets are prone to losing details in the continuous downsampling process, and the blurred boundary makes the segmentation accuracy less satisfactory. Therefore, the segmentation of calcified plaques is a very challenging task. To address the above problems, in this paper, we design a framework named LPE-UNet, which adopts an encoder-decoder structure similar to UNet. The framework includes two powerful modules named the low-rank perception enhancement module and the noise filtering module. The low-rank perception enhancement module extracts multi-scale context features by increasing the receptive field size to aid target detection and then uses an attention mechanism to filter out redundant features. The noise filtering module suppresses noise interference in shallow features to high-level features in the process of multi-scale feature fusion. It computes a pixel-wise weight map of low-level features and filters out useless and harmful information. To alleviate the problem of class imbalance caused by small-sized lesions, we use a weighted cross-entropy loss function and Dice loss to perform mixed supervised training on the network. The proposed method was evaluated on the calcified plaque segmentation dataset, achieving a high F1 score of 0.941, IoU of 0.895, and Dice of 0.944. This result verifies the effectiveness and superiority of our approach for accurately segmenting calcified plaques. As there is currently no authoritative publicly available calcified plaque segmentation dataset, we have constructed a new dataset for coronary artery calcified plaque segmentation (Calcified Plaque Segmentation Dataset, CPS Dataset).  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
2079-9292  
入藏号:  
WOS:001017032100001  
Indexed Date:  
2023-07-08  
Institution Address:  
Shandong Univ, Sch Software, Jinan 250101, Peoples R China  
Institution:

### 1001. EGD-Net: Edge-guided and differential attention network for surface

摘要: Establishing an automated defect detection system is a critical task for industrial production, but the current defect detection system still faces great challenges, especially for defects with blurred edges and weak defects in complex backgrounds. To solve these problems, we propose an edge-guided and differential attention network (EGD-Net), which can highlight the defect areas by strengthening edge information and effectively eliminating background clutter. In the proposed network, the multi-scale features are first extracted. Then, a specially designed edge prediction module is used to extract defect edge information from shallow layers. Three multi -scale feature fusion modules are employed to fuse context information in the deep layers. Following this, the edge fusion module is constructed to complement the edge information and context information for better guiding defect segmentation. Finally, a differential attention module (DAM) is designed to perform top-down attention and produce the final prediction results. The DAM can effectively eliminate clutter in the back-ground area caused by connecting the edge features. In the experiment, we collected a packaging box dataset with complex background patterns from the practical industrial field to verify the proposed model performance. Moreover, four public datasets were also employed to validate the model. Experimental results (mIoU/mPA) (DAGM2007: 85.53%/88.19%, CrackForest: 87.58%/91.32%, AITEX defect: 78.31%/82.29%, MT defect: 77.08%/81.19%, box defect 94.39%/96.51%:) show that our proposed method outperforms other state-of-the-art methods, especially for the detection of complex background defects.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
10  
ISSN:  
2467-964X  
eISSN:  
2452-414X  
入藏号:  
WOS:000867374400001  
Indexed Date:  
2022-10-24  
Institution Address:  
Xian Univ Technol, Sch Mech & Precis Instrument Engn, Xian 710048, Peoples R China  
Institution Address:  
Xian Univ Technol, Dept Informat Sci, Xian 710048, Peoples R China  
Institution Address:  
Xian Univ Technol, Sch Comp Sci & Engn, Xian 710048, Peoples R China  
Institution:  
Institution:  
Institution:

### 1002. A scale and region-enhanced decoding network for nuclei classification

摘要: Accurate classification of nuclei in histology images is essential for clinical diagnosis, prognosis, and therapeutic response prediction of cancer. However, this is still a challenging task due to (1) nuclei exhibiting a high level of heterogeneity within different types and (2) large intra-class variability including complex morphology and large variations of scale. To solve these problems, we propose a novel scale and region-enhanced decoding network based on the traditional U-shape structure for nuclei classification. We employ a nuclei detection head as region enhancement module in the decoding branch, which can enhance the nuclear regional information by locating the approximate bounding regions and provide more distinguish information for producing better feature maps of subsequent classification. Then, we propose a scale-aware feature fusion module, which fuses stage-wise feature maps generated from the decoder branch, to effectively learn multi-scale features. Finally, we utilize a scale attention module to calibrate the features and adapt to the most suitable scale in the hybrid multi -scale feature maps. In comparison with several state-of-the-art methods on two publicly available colonic cancer nuclei classification datasets, namely ConSep and Lizard, the proposed method obtains the highest accuracy of 0.860 and 0.927, respectively. It also achieves the highest accuracy of 0.838 on the PanNuke dataset collected from different tissues at different magnitudes. The independent validation on two subsets of the Lizard dataset indicates the proposed method obtains the highest accuracy. In conclusion, the proposed method can greatly improve classification performance, particularly for challenging nuclei with complex contexts and large-scale variations.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:000931780800001  
Indexed Date:  
2023-03-05  
Institution Address:  
Univ South China, Sch Comp, Hengyang 421001, Peoples R China  
Institution:

### 1003. Online surface defect segmentation on aluminum strip production line

摘要: A significant challenge in automated defect inspection (ADI) of aluminum strip surfaces is improving segmentation speed to satisfy the online inspection requirements of the production line while maintaining the precision of defect identification. This study proposes a lightweight and efficient defect segmentation model that can be applied in aluminum processing enterprises for fast and precise segmentation of aluminum strip surface defects. A novel fusion attention (FA) mechanism is first established to enhance the focus on critical characteristics along the spatial and channel dimensions. This mechanism adopts continuous dilated convolutions with appropriate dilation rates to effectively increase the range of the receptive field and improve defect localization accuracy. Subsequently, a lightweight MobileViTv2 with an embedded FA mechanism is employed as a multi-scale feature extractor to learn comprehensive representations from defect images. Next, a novel feature fusion method, named large-scale feature pyramid network (LSFPN), is introduced to enhance the focus on details within large-scale features. LSFPN establishes four progressively shallower top-down pathways with fast normalized fusion weights and incorporates lightweight aggregation nodes based on the MoblieNetv2 block. Surface images of straightened aluminum strips with five universal defects were collected, whereby a new dataset was established. The experimental outcomes demonstrate the proposed model outperforms other state-of-the-art techniques synthetically, achieving a mean Intersection over Union (mIoU) of 87.01%, a segmentation speed of 61.67 fps, and a model size of 16.23 MB. This model may serve as a valuable theoretical foundation for the online segmentation of aluminum strip surface defects in embedded devices.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001070648400001  
Indexed Date:  
2023-10-05  
Institution Address:  
Cent South Univ, Light Alloy Res Inst, Changsha 410083, Hunan, Peoples R China  
Institution Address:  
Guangxi Liuzhou Yinhai Aluminum Co Ltd, Liuzhou 545006, Guangxi, Peoples R China  
Institution:  
Institution:  
Guangxi Liuzhou Yinhai Aluminum Co Ltd

### 1004. Aerial military target detection algorithm based on multi-feature cross

摘要: The precise detection of military targets under complex conditions is a key factor to enhance the ability of war situation generation and prediction. The current technology can not overcome the problems of smoke and occlusion interference,target height change,and different scales in aerial video. In this paper,a multi feature cross fusion and cross layer cascade aerial military target detection algorithm (YOLOv5-MFLC) is proposed. Firstly,aiming at the high confidentiality of the military targets and the shortage of battlefield aerial image resources,a real scene based aerial military target dataset is constructed,and the methods of random splicing and random extraction embedding are used for data enhancement in order to improve the diversity and generalization of targets. Secondly,aiming at the problem of complex background interference,a multi feature cross fusion attention mechanism is constructed to enhance the available information of target features. Finally,for the multi-scale problem of targets in aerial images,a cross layer cascaded multi-scale feature fusion pyramid is designed to improve the detection accuracy of cross scale targets. The experimental results show that,comparing with the existing advanced detection models, the detection accuracy of the algorithm in this paper has been greatly improved. The average accuracy of the algorithm can reach 81.0%,which is 5.2% higher than the original network. In particular,it has reached 55.9% in the smaller target category " person",which is 9.4% higher. And the experimental results further show the usefulness of the improved algorithm for small target detection. At the same time,the detection rate of this algorithm can reach 56 frame/s,which can effectively achieve accurate and fast detection of battlefield targets,and has certain experience value for guiding complex modern wars.  
摘要:  
复杂条件下特殊目标的精确检测是增强特定场景态势生成和预测能力的关键因素。目前的技术不能克服航拍视频中出现的烟雾和遮挡干扰、目标高度变化、尺度不一等问题。因此,提出一个多特征交叉融合及跨层级联的航拍特殊目标检测算法(YOLOv5-MFLC) 。针对实际特殊目标保密性高、航拍图像资源匮乏的问题,构建了一个基于真实场景的航拍特殊目标数据集,并采用随机拼接和随机提取嵌入的方法进行数据增强以提高目标多样性和泛化性;针对复杂背景干扰问题,构建了多特征交叉融合注意力机制,增强了目标特征的可用信息;针对航拍图像中目标多尺度问题,设计了跨层级联多尺度特征融合金字塔,提高了跨尺度目标的检测准确率。实验结果表明,与现有的先进检测模型相比,所提算法的检测准确率有较大提升,算法平均准确率可达到81.0%,相比于原始网络提升了5.2%,特别是,在更小的目标类别person中达到了55.9%,提升了9.4%,进一步表明了所提改进算法对小目标检测的有用性。同时,所提算法的检测速率可以达56 frame/s,能够有效地实现实际复杂场景特殊目标的准确、快速检测,对特殊目标的识别具有一定的指导意义。  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1000-2758  
入藏号:  
CSCD:7619893  
Indexed Date:  
2024-03-02  
Institution Address:  
西安工业大学计算机科学与工程学院, 西安, 陕西 710021, 中国  
Institution Address:  
西安工业大学兵器科学与技术学院, 西安, 陕西 710021, 中国  
Institution Address:  
西安工业大学机电工程学院, 西安, 陕西 710021, 中国  
Institution Address:  
School of Computer Science and Engineering,Xi'an University of Technology, Xi'an, Shaanxi 710021, China  
Institution Address:  
School of Weapon Science and Technology,Xi'an University of Technology, Xi'an, Shaanxi 710021, China  
Institution Address:  
School of Mechanical and Electrical Engineering,Xi'an University of Technology, Xi'an, Shaanxi 710021, China  
Institution:  
西安工业大学计算机科学与工程学院  
Institution:  
西安工业大学兵器科学与技术学院  
Institution:  
西安工业大学机电工程学院  
Institution:  
School of Computer Science and Engineering,Xi'an University of Technology  
Institution:  
School of Weapon Science and Technology,Xi'an University of Technology  
Institution:  
School of Mechanical and Electrical Engineering,Xi'an University of Technology

### 1005. Self-supervised multi-magnification feature enhancement for segmentation

摘要: Accurate segmentation of hepatocellular carcinoma regions is a critical step in surgical evaluation of whole slide pathological images. Recently, emerging multi-magnification learning-based methods have shown promise with evaluation of whole slide images. However, traditional multi-magnification learning segmentation models focus on uniformity and fail to effectively exploit different magnification information. The purpose of this study is to develop a novel multi-magnification learning segmentation model by effectively utilizing diverse magnification information. Therefore, we developed a novel multi-magnification feature self-enhancement network to extract magnification-specific features from multi-magnification images via self-supervised learning without extra supervision. Specifically, the proposed network enhances feature information via a super-resolution module based on self-supervision to transfer the representative information into a segmentation encoder, which takes full advantage of the pyramid storage characteristic of whole slide images. Moreover, a multi-scale feature fusion module based on the attention mechanism is used to fuse pretrained multi-scale features through a gated unit block, which is designed to fuse features from generation task into and segmentation task. We evaluated this method using the hepatocellular carcinoma data set from The Cancer Genome Atlas with cross-validation and achieved a 0.829 Dice similarity coefficient, nearly a 3% improvement over state-of-the-art models used for segmentation. To the best of our knowledge, this is the first study of the application of super-resolution as a selfenhancement network for histopathological image segmentation via self-supervised pretext learning. The code will be available at https://github.com/SH-Diao123/Self-Supervised-Multi-magnification-Segmentation.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
2  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001225518500001  
Indexed Date:  
2024-05-25  
Institution Address:  
Chinese Acad Sci, Shenzhen Inst Adv Technol, Shenzhen 518055, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Shenzhen Coll Adv Technol, Shenzhen 518055, Peoples R China  
Institution Address:  
Univ Texas MD Anderson Canc Ctr, Dept Imaging Phys, Div Diagnost Imaging, Houston, TX 77030 USA  
Institution:  
Institution:  
Institution:

### 1006. MA-YOLO: A Method for Detecting Surface Defects of Aluminum Profiles

摘要: Aluminum Profiles (APs) are aluminum materials obtained by hot melting and extruding aluminum rods. It has the characteristics of low cost, strong plasticity, easy processing, and recyclability, and therefore plays an important role in industrial production. However, defects such as Non-Conductive (NC), Scratch, Orange Peel (OP), and Dirty Point (DP) often occur during the production and processing of APs, which can seriously affect the quality of APs. In addition, surface defects of APs also have problems such as fuzzy regional definition, large-scale variation, imbalance of aspect ratio, and high inter-class defect similarity, making defect detection more challenging. To solve these problems, this paper proposes an attention-guided object detection algorithm called MA-YOLO, specifically for Surface Defect Detection (SDD) of APs. The algorithm is based on YOLOv5s. Firstly, the K-Means++ clustering algorithm is used to optimize the anchor boxes, which alleviates the problem of aspect ratio imbalance. Secondly, by improving the multi-scale Feature Fusion Network (FFN), the detection performance of the model to detect the defects with unbalanced aspect ratio is improved, and the adaptability of the model to defects of different scales is enhanced. Finally, a novel Max Pooling Average Pooling (MA) attention module is proposed to improve the overall detection performance of the model, especially for small-scale defects. Experimental results on the aluminum profile surface defect dataset show that MA-YOLO has better detection performance and superiority than the current mainstream object detection algorithms, and compared with the baseline YOLOv5s, the mAP(50) and F1 score are increased by 2.9% and 2.2%, respectively, while keeping the model lightweight. This indicates that MA-YOLO has broad application prospects in the surface defect detection of APs.  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
11  
被引频次合计:  
12  
ISSN:  
2169-3536  
入藏号:  
WOS:001033514000001  
Indexed Date:  
2023-08-08  
Institution Address:  
Xijing Univ, Sch Elect Informat, Xian 710123, Shaanxi, Peoples R China  
Institution Address:  
Shaanxi Key Lab Integrated & Intelligent Nav, Xian 710068, Shaanxi, Peoples R China  
Institution Address:  
Xijing Univ, Xian Key Lab High Precis Ind Intelligent Vis Measu, Xian 710123, Shaanxi, Peoples R China  
Institution Address:  
Xijing Univ, Grad Off, Xian 710123, Shaanxi, Peoples R China  
Institution Address:  
Xijing Univ, Sch Comp Sci, Xian 710123, Shaanxi, Peoples R China  
Institution:  
Institution:  
Shaanxi Key Lab Integrated & Intelligent Nav  
Institution:  
Institution:  
Institution:

### 1007. CBAM plus ASFF-YOLOXs: An improved YOLOXs for guiding agronomic

摘要: To the optimal time to conduct farming operations in the traditional agricultural production process mainly depends on human observation and planting experience, which is time-consuming and laborious, and makes it easy to miss the best agricultural operation opportunities. In this study, our main objective is to accurately detect the key growth stages of lettuce to guide the timely implementation of corresponding agricultural operations. Firstly, the dataset was collected for the growth stage with important agricultural operations in the growth process of multi-variety lettuce, to lay the data foundation for the construction of the model. Secondly, considering the difference in plant growth, we compared many methods and selected the optimal modeling method YOLOXs to identify the key growth stages of multi-variety lettuce (mAP = 98.75 %). Finally, to ensure the applicability of the detection model in complex agricultural scenes, we tried to improve the effect of YOLOXs by three attention mechanisms and one multi-scale feature fusion method, and proposed a new method CBAM + ASFF-YOLOXs (mAP = 99.04 %). The results showed that this method is expected to replace human eye observation and experience in planting, to provide accurate technical feedback on relevant agricultural operation time, and to provide technical support for the unmanned operation of agriculture. At the same time, the limi-tations, challenges, and prospects of this method are discussed.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
9  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:000900019500001  
Indexed Date:  
2023-01-13  
Institution Address:  
China Agr Univ, Natl Innovat Ctr Digital Fishery, Beijing, Peoples R China  
Institution Address:  
China Agr Univ, Key Lab Smart Farming Technol Aquat Anim & Livesto, Minist Agr & Rural Affairs, Beijing 100083, Peoples R China  
Institution Address:  
China Agr Univ, Beijing Engn & Technol Res Ctr Internet Things Agr, Beijing 100083, Peoples R China  
Institution Address:  
China Agr Univ, Coll Informat & Elect Engn, Beijing 100083, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 1008. Improved YOLOv8 algorithms for small object detection in aerial imagery

摘要: In drone aerial target detection tasks, a high proportion of small targets and complex backgrounds lead to false positives and missed detections, resulting in low detection accuracy. To improve the accuracy of the detection of small targets, this study proposes two improved models based on YOLOv8s, named IMCMD\_YOLOv8\_small and IMCMD\_YOLOv8\_large. Each model accommodates different application scenarios. First, the network structure was optimized by removing the backbone P5 layer used to detect large targets and merging the P4, P3, and P2 layers, which are better suited for detecting medium and small targets; P3 and P2 serve as detection heads to focus more on small targets. Subsequently, the coordinate attention mechanism is integrated into the backbone's C2f, to create a C2f\_CA module that enhances the model' s on key information and secures a richer flow of gradient information. Subsequently, a multiscale attention feature fusion module was designed to merge the shallow and deep features. Finally, a Dynamic Head introduced to unify the perception of scale, space, and tasks, further enhancing the detection capability for small targets. Experimental results on the VisDrone2019 dataset demonstrated that, compared YOLOv8s, IMCMD\_YOLOv8\_small achieved improvements of 7.7% and 5.1% in mAP@0.5 and mAP@0.5:0.95, respectively, with a 73.0% reduction in the parameter count. The IMCMD\_YOLOv8\_large model showed even more significant improvements in these metrics, reaching 10.8% and 7.3%, respectively, with a 47.7% reduction in the parameter count, displaying superior performance in small target detection tasks. improved models not only enhanced the detection accuracy but also achieved model lightweighting, thereby proving the effectiveness of the improvement strategies and showcasing superior performance compared other classic models.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1319-1578  
eISSN:  
2213-1248  
入藏号:  
WOS:001264427600001  
Indexed Date:  
2024-07-16  
Institution Address:  
Kaili Univ, Sch Big Data Engn, Kaili 556000, Peoples R China  
Institution:

### 1009. High-precision target ranging in complex orchard scenes by utilizing

摘要: The automation of orchard production is increasingly relying on robotics, driven by the advancements in artificial intelligence technology. However, accurately comprehending semantic information and precisely locating various targets within orchard environments remain challenges. Current research often relies on expensive multisensor fusion techniques or vision-only approaches that yield inadequate segmentation outcomes for perceiving orchard surroundings. To address these issues, this article proposes a novel approach for target ranging in complex orchard scenes, leveraging semantic segmentation results. The article introduces the MsFF-Segformer model, which employs multi-scale feature fusion to generate high-precision semantic segmentation images. The model incorporates the MiT-B0 encoder, which utilizes a pure attention mechanism, and the MsFF decoder, specifically designed for multi-scale feature fusion. The MsFF decoder includes the AFAM module to effectively align features of adjacent scales. Additionally, the channel attention module and depth separable convolution module are introduced to reduce model parameter size and obtain feature vectors with rich semantic levels, enhancing the segmentation performance of multi-scale targets in orchards. Based on the accurate semantic segmentation outcomes in orchard environments, this study introduces a novel approach named TPDMR that integrates binocular vision to estimate the distances of various objects within orchards. Firstly, the process involves matching the semantic category matrix with the depth information matrix. Subsequently, the depth information array that represents the target category is obtained, and any invalid depth information is filtered out. Finally, the average depth of the target is calculated. Evaluation of the MsFF-Segformer model on a self-made orchard dataset demonstrates superior performance compared to U-net and other models, achieving a Mean Intersection over Union (MIoU) of 86.52 % and a Mean Pixel Accuracy (MPA) of 94.05 %. The parameters and prediction time for a single frame are 15.1 M and 0.019 s, respectively. These values are significantly lower than those of U-net, Deeplabv3+, and Hrnet models, with reductions of 84.1 %, 32.5 %, 5.9 % and 69.4 %, 59.7 %, 64.2 % respectively. The TPDMR method demonstrates a high level of accuracy and stability in target ranging, with a ranging error of less than 6 % across all targets. Furthermore, the overall algorithm runtime is estimated to be approximately 0.8 s, indicating efficient performance.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:001126550500001  
Indexed Date:  
2023-12-30  
Institution Address:  
Nanjing Agr Univ, Coll Engn, Nanjing 210031, Peoples R China  
Institution:

### 1010. An efficient multi-task convolutional neural network for dairy farm

摘要: Real-time and accurate detection of multiple types of targets and obstacles in dairy barns is a necessary function for autonomous pushing robots. To improve the efficiency of target recognition and to reduce the path extraction error of the pushing robot, on the basis of the high accuracy perception of every pixel collected with an embedded AI computer, a multi-task learning based dairy barn multi-type target recognition model Ghost CBAM Segmentation-Multi-task (GCS-MUL) was proposed, which could recognize dairy cows, obstacles and road targets in real-time and efficiently. Firstly, in order to enhance the ability to extract key features from the targets, the proposed model intergrades the Convolutional Block Attention Module (CBAM), a self-designed light-weight target feature extraction network Ghost CBAM Network (GCNet) as the backbone of the whole model. Secondly, to improve the model multi-scale feature fusion, Path Aggregation Network (PAN) and Feature Pyramid Network (FPN) structures with the GhostConv module were used in neck net. Finally, for real-time semantic segmentation dairy farms multiple targets, a Segmentation Head (Seg Head), which is composed of the Receptive Field Block (RFB), Pyramid Pooling Module (PPM) and Feature Fusion Module (FFM), was introduced. Experimental results showed that the mAP@0.5 (mean average precision IoU = 0.5) of the dairy farm target reached 94.86%. Compared to the YOLOv5 model, the precision and recall was improved by 7.47% and 6.85%, respectively. In comparison to the YOLOv7 model, the precision was improved by 5.1%. Furthermore, when compared to the SSD model, the proposed model have reduced the number of model parameters by 92.43%, and its average detection time was reduced by 84.37 ms, which is ideal for meeting the real-time target recognition requirements. The average detection time of the model is 66.43 ms, making it more suitable for deployment in embedded devices. Compared with Ghost CBAM-Detection (GC-Detect) without the introduction of the Seg Head, the precision, recall and mAP@0.5 was improved by 4.49%, 4.92% and 6.58%, respectively. The research results can provide accurate algorithms for real-time and efficient identification of dairy farm targets for pushing robots, and provide more effective road and environmental scene segmentation methods for autonomous walking.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
6  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:001024972600001  
Indexed Date:  
2023-07-21  
Institution Address:  
Shandong Agr Univ, Coll Mech & Elect Engn, Tai An 271018, Peoples R China  
Institution Address:  
Shandong Prov Engn Lab Agr Equipment Intelligence, Tai An 271018, Peoples R China  
Institution Address:  
Shandong Prov Key Lab Hort Machineries & Equipment, Tai An 271018, Peoples R China  
Institution Address:  
Shandong Agr Univ, Coll Life Sci, Tai An 271018, Peoples R China  
Institution Address:  
Chinese Acad Agr Mechanizat Sci, Beijing 100083, Peoples R China  
Institution Address:  
Shandong Taikai High Voltage Switchgear Co Ltd, Tai An 271000, Peoples R China  
Institution Address:  
Shandong Agr Univ, Coll Anim Sci & Technol, Tai An 271018, Peoples R China  
Institution:  
Institution:  
Shandong Prov Engn Lab Agr Equipment Intelligence  
Institution:  
Shandong Prov Key Lab Hort Machineries & Equipment  
Institution:  
Institution:  
Chinese Acad Agr Mechanizat Sci  
Institution:  
Shandong Taikai High Voltage Switchgear Co Ltd  
Institution:

### 1011. YOFIR: High precise infrared object detection algorithm based on YOLO

摘要: Infrared images often suffer from issues such as blurriness and unclear object boundaries, and existing object detection algorithms are developed based on visible light images, which makes infrared object detection more challenging. Therefore, this paper proposes an infrared image enhancement method and an infrared object detection algorithm based on YOLO and FasterNet, named YOFIR. Specifically, we apply CHALE, Auto Gamma, histogram equalization, and bilateral filtering to process images individually, then fuse the results with different weights to address the poor imaging quality of infrared images. Moreover, we utilize the FasterNet network for multi-scale feature extraction to adapt to low-resolution infrared images. We also reduce model parameters through GSConv and propose a novel Efficient Multi-Scale Group Convolution module, EMSGC, which enhances feature fusion by processing feature maps from different channels, effectively improving detection accuracy. Finally, the DyHead Block is incorporated into the head to enhance the capability of infrared object detection. Experimental results on the HIT-UAV infrared remote sensing dataset show that the proposed algorithm achieves a 4% improvement in mAP0.5 compared to YOLOv8. Moreover, on the FLIR dataset, the algorithm shows a 1.6% improvement in mAP 0 . 95 over YOLOv8, with significant advantages in terms of model parameters and FLOPs.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1350-4495  
eISSN:  
1879-0275  
入藏号:  
WOS:001364429400001  
Indexed Date:  
2024-12-03  
Institution Address:  
Shanghai Univ Elect Power, Coll Comp Sci & Technol, Shanghai 201306, Peoples R China  
Institution Address:  
State Grid Shanghai Elect Power Co Ltd, Shanghai 200122, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Sch Cyber Sci & Engn, Shanghai 200240, Peoples R China  
Institution:  
Institution:  
State Grid Shanghai Elect Power Co Ltd  
Institution:

### 1012. ECFNet: Efficient cross-layer fusion network for real time RGB-Thermal

摘要: RGB-Thermal semantic segmentation is widely used in various perception scenarios and has made significant progress. However, many existing methods tend to overlook the critical challenge of striking a balance between speed and accuracy. In response to this, we introduce ECFNet, an efficient model tailored for real-time RGBThermal semantic segmentation, aiming to achieve a balance between speed and accuracy to some extent. Specifically, we enhance feature fusion by incorporating the Asymmetric Cross-layer Self-Attention (ACSA) module, enabling the amalgamation of feature maps across diverse intermediate layers. Additionally, we introduce the Light Effective Spatial Semantic Fusion (LESSF) module to merge feature maps from the final layer. To fully exploit the latent multi-modal feature information, we introduce a Multi-branch Cascade Decoder (MCD) composed of six Hybrid Attention Module (HAM) blocks, it is used to aggregate multi-scale feature maps. We validated our approach on three publicly available benchmark datasets including MFNet, PST900, and FMB. Our method shows effectiveness and achieves a better balance between speed and accuracy. ECFNet achieved 56.2% mIoU on the MFNet dataset and 62.4 FPS on a single NVIDIA GeForce GTX 1080Ti GPU. The code and results are available at https://github .com /WangJoyu /ECFNet.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1051-2004  
eISSN:  
1095-4333  
入藏号:  
WOS:001333795000001  
Indexed Date:  
2024-10-22  
Institution Address:  
China Univ Min & Technol, Sch Informat & Control Engn, Xuzhou 221116, Peoples R China  
Institution:

### 1013. Improving PCB defect detection using selective feature attention and

摘要: - Due to the ongoing miniaturization of electronic products and the use of miniature printed circuit boards (PCBs), existing AI-based defect detection methods have exhibited poor performance in detecting tiny PCB defects. This issue can potentially compromise safety, degrade manufacturing quality, and increase production costs. To tackle this problem, we propose two novel techniques for PCB defect detection, namely Selective Feature Attention (SF attention) and Pixel Shuffle Pyramid (PSPyramid). SF attention identifies important features from a pyramid feature map to fuse the semantic and spatial information, while PSPyramid effectively fuses semantic features to detect various types of defects on PCBs, especially tiny defects. Moreover, a customized training strategy, specifically for PCB defect detection, is devised. To evaluate the performance of our proposed algorithms, extensive experiments have been conducted on two well-known PCB datasets containing tiny defects: the DeepPCB and TDD datasets. Our proposed non-referential method achieves performance comparable to existing referential methods on the DeepPCB dataset, making it more feasible for industrial applications. Compared to state-of-the-art methods, our method reduces the error by 16%, in terms of AP50, on the TDD dataset. The experimental results demonstrate the effectiveness of our proposed method in improving the quality assurance process for PCBs in the electronics industry.  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
2590-1230  
入藏号:  
WOS:001208489600001  
Indexed Date:  
2024-05-04  
Institution Address:  
Hong Kong Polytech Univ PolyU, Hong Kong, Peoples R China  
Institution Address:  
Ctr Adv Reliabil & Safety Ltd CAiRS, Hong Kong, Peoples R China  
Institution:  
Institution:  
Ctr Adv Reliabil & Safety Ltd CAiRS

### 1014. Image Reconstruction Based on Fused Features and Perceptual Loss

摘要: Compressive sensing (CS) technology is introduced into space optical remote sensing image acquisition stage, which could make wireless image sensor network node quickly and accurately obtain images in the case of two constraints of limited battery power and expensive sensor costs. On this basis, in order to further improve the quality of CS image reconstruction, we propose fused features and perceptual loss encoder-decoder residual network (FFPL-EDRNet) for image reconstruction. FFPL-EDRNet consists of a convolution layer and a reconstruction network. We train FFPL-EDRNet end-to-end, thus greatly simplifying the pre-processing and post-processing process and eliminating the block effect of reconstructed images. The reconstruction network is based on residual network, which introduces multi-scale feature extraction, multi-scale feature combination and multi-level feature combination. Feature fusion integrates low-level information with high-level information to reduce reconstruction error. The perceptual loss function based on pretrained InceptionV3 uses the weighted mean square error to define the loss value between the reconstructed image feature and the label image feature, which makes the reconstructed image more semantically similar to label image. In the measurement procedure, we use convolution to achieve block compression measurement, so as to obtain full image measurements. For image reconstruction, we firstly use a deconvolution layer to initially reconstruct the image and then use the residual network to refine the initial reconstructed image. The experimental results show that: in the case of measurement rates (MRs) of 0.25, 0.10, 0.04 and 0.01, the peak signal-to-noise ratio (PSNR) = 27.502, 26.804, 24.593, 21.359 and structural similarity (SSIM) = 0.842, 0.816, 0.720, 0.568 of the reconstructed images obtained by FFPL-EDRNet. Therefore, Our FFPL-EDRNet could enhance the quality of image reconstruction.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
2169-3536  
入藏号:  
WOS:000638398500001  
Indexed Date:  
2021-04-18  
Institution Address:  
Chinese Acad Sci, Changchun Inst Opt Fine Mech & Phys, Changchun 130033, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Daheng Coll, Beijing 100039, Peoples R China  
Institution Address:  
Chinese Acad Sci, Key Lab Space Based Dynam & Rapid Opt Imaging Tec, Changchun 130033, Peoples R China  
Institution:  
Institution:  
Institution:

### 1015. Automatic pixel-level crack detection with multi-scale feature fusion

摘要: Cracks are common defects in slab tracks, which can grow and expand over time, leading to a deterioration of the mechanical properties of slab tracks and shortening service life. Therefore, it is essential to accurately detect and repair cracks before they impact services. This study developed a systematic pixel-level crack segmentation-quantification method suited for nighttime detection of slab tracks. To be specific, slab track crack network II, a pixel-level segmentation network that aggregates multi-scale information was proposed to extract the morphology of slab track cracks, and then their widths were calculated by an alternative quantification method proposed in the paper. The model performs best when the initial learning rate is 0.0001, with intersection over unions (IOUs) 84.94% and 83.84% observed on the training set and validation set, respectively. In the test set, the IOU value is 81.07%, higher than that derived from similar segmentation algorithms, indicating higher robustness and better generalization of the network architecture. In addition, the average errors in predicting crack widths resulting from the proposed method are 0.13 and 0.12 mm, compared to the results measured by a vernier caliper and a 3D scanner, respectively. The proposed pixel-level segmentation-quantification system provides a new method and theoretical support for slab track maintenance and repair.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
50  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
50  
ISSN:  
1093-9687  
eISSN:  
1467-8667  
入藏号:  
WOS:000939622800001  
Indexed Date:  
2023-03-17  
Institution Address:  
Southwest Jiaotong Univ, Sch Civil Engn, Chengdu, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, MOE Key Lab High Speed Railway Engn, Chengdu, Peoples R China  
Institution Address:  
China Railway Soc, Beijing, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, Sch Civil Engn, Chengdu 610031, Peoples R China  
Institution:  
Institution:  
Institution:  
China Railway Soc  
Institution:

### 1016. Coronary artery segmentation in CCTA images based on multi-scale feature

摘要: BACKGROUND: Coronary artery segmentation is a prerequisite in computer-aided diagnosis of Coronary Artery Disease (CAD). However, segmentation of coronary arteries in Coronary Computed Tomography Angiography (CCTA) images faces several challenges. The current segmentation approaches are unable to effectively address these challenges and existing problems such as the need for manual interaction or low segmentation accuracy. OBJECTIVE: A Multi-scale Feature Learning and Rectification (MFLR) network is proposed to tackle the challenges and achieve automatic and accurate segmentation of coronary arteries. METHODS: The MFLR network introduces a multi-scale feature extraction module in the encoder to effectively capture contextual information under different receptive fields. In the decoder, a feature correction and fusion module is proposed, which employs high-level features containing multi-scale information to correct and guide low-level features, achieving fusion between the two-level features to further improve segmentation performance. RESULTS: The MFLR network achieved the best performance on the dice similarity coefficient, Jaccard index, Recall, F1-score, and 95% Hausdorff distance, for both in-house and public datasets. CONCLUSION: Experimental results demonstrate the superiority and good generalization ability of the MFLR approach. This study contributes to the accurate diagnosis and treatment of CAD, and it also informs other segmentation applications in medicine.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0895-3996  
eISSN:  
1095-9114  
入藏号:  
WOS:001315146200005  
PubMed ID:  
38943423  
Indexed Date:  
2024-09-23  
Institution Address:  
Northeastern Univ, Coll Med & Biol Informat Engn, Shenyang, Peoples R China  
Institution Address:  
Northeastern Univ, Software Coll, Shenyang, Peoples R China  
Institution Address:  
Gen Hosp North Theater Command, Dept Radiol, Shenyang, Liaoning, Peoples R China  
Institution Address:  
Minist Educ, Key Lab Med Image Comp, Shenyang, Peoples R China  
Institution Address:  
Minist Educ, Engn Res Ctr Med Imaging & Intelligent Anal, Shenyang, Peoples R China  
Institution Address:  
Macquarie Univ, Fac Med Hlth & Human Sci, Macquarie Med Sch, Sydney, Australia  
Institution:  
Institution:  
Institution:  
Gen Hosp North Theater Command  
Institution:  
Minist Educ  
Institution:  
Minist Educ  
Institution:

### 1017. CM-YOLO: Context Modulated Representation Learning for Ship Detection

摘要: Ship detection is essential for both military and civilian applications. Existing ship detection methods focus on prominent offshore ships, paying less attention to complex nearshore ships, which are easily confused with the intricate background. Utilizing contextual information, such as location and shape, can enhance ship detection and classification in complex environments. In this article, we propose a context modulated representation learning-based detection method termed as CM-YOLO. It adopts the classical detector design framework, which includes the backbone, neck, and head. The input image is sequentially processed through these components to obtain the detection results. Our method specifically optimizes ship detection in complex scenarios. To achieve this, we propose a dual path context enhancement neck (DCEN) to extract contextual information for ship detection. The neck builds on the path augmentation feature pyramid network with the proposed dual path context enhancement (DCE) module, which is designed to enhance feature representations by incorporating high-level semantic information. It captures long-range dependencies across both channel and spatial dimensions while suppressing irrelevant features. Additionally, to enhance the scale-aware capability of the head for detecting multiscale ships in complex environments, we introduce the multicontext boosted (MCB) detection head. The MCB can flexibly adjust the receptive field and extracts relevant context for ships of various scales using multiple large-kernel convolutions. We conduct experiments on three commonly used ship datasets: Seaships7000, ShipRSImageNet, DIOR-ship, and HRSC2016. Experiment results demonstrate that CM-YOLO achieves excellent performance compared with other leading ship detection methods.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001425930700014  
Indexed Date:  
2025-02-27  
Institution Address:  
Northwestern Polytech Univ, Sch Elect & Informat, Xian 710072, Peoples R China  
Institution Address:  
Dalian Maritime Univ, Coll Marine Elect Engn, Dalian 116026, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Unmanned Syst Res Inst, Xian 710072, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Sch Astronaut, Xian 710072, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 1018. Dual-path multi-scale attention residual network for fault diagnosis of

摘要: Rolling bearing faults inevitably occur during the long-term continuous operation of rotating machinery. Therefore, fault diagnosis is greatly important for ensuring the normal and safe operation of rolling bearings. However, the complexity and diversity of working conditions of rolling bearings present a significant challenge in extracting fault characteristics accurately, which further affects the ultimate fault diagnosis results. In this article, we propose a new model, called dual-path multi-scale attention residual network (DPMARN), for diagnosing bearing faults under complex operating conditions. DPMARN can effectively capture the feature-feature correlation information at different scales, which is more beneficial for fusing fault features at different scales to improve the model's performance. The main contributions of this work are summarized as follows: (1) the designed dual-path network model which incorporates parallel multi-scale branches of convolutional kernels and serially connects skip-layer multi-scale branches can integrate both low-frequency and high-frequency information and enhance the multi-scale feature extraction and complex data representation abilities. (2) The squeeze-and-excitation attention mechanism is embedded into the residual blocks to improve the ability of learning feature correlations and utilizing feature information effectively, which is helpful for extracting important fault characteristics. Extensive experiments conducted on two public bearing datasets demonstrate the superior performance of the DPMARN model for addressing the complex fault diagnosis problem. These results indicate that our proposed approach provides an effective solution for fault diagnosis of rolling bearings under complex operating conditions.  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001215215200001  
Indexed Date:  
2024-05-11  
Institution Address:  
Lanzhou Univ Technol, Sch Mech & Elect Engn, Lanzhou, Peoples R China  
Institution:

### 1019. Underwater image enhancement method based on the generative adversarial

摘要: Aiming at the problems of color distortion, nonuniform illumination, and low contrast caused by degradation of underwater images, an underwater image enhancement method (MSFF-GAN) based on generative adversarial network was proposed. A multiscale featured fusion generator is designed, which improves the ability to use different scale features of the model and ensures that the generated image retains more detailed information. The residual dense module is constructed to solve the problem of generator characteristics extracted slower. In the discriminator, to achieve the extraction of local image features, the output matrix is discriminating so that the generated image is closer to the real image. Compared with the existing underwater image enhancement methods qualitatively and quantitatively, the proposed method has better enhancement effect on EUVP and RUIE datasets. The proposed method is superior to the contrast method of three evaluation indexes: PSNR, SSIM, and UIQM. ? 2021 SPIE and IS&T [DOI: 10.1117/1.JEI.30.1.013009]  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:000634148400006  
Indexed Date:  
2021-04-18  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao, Peoples R China  
Institution Address:  
Shandong Univ Sci & Technol, Shandong Prov Key Lab Wisdom Mine Informat Techno, Qingdao, Peoples R China  
Institution:  
Institution:

### 1020. An efficient model for metal surface defect detection based on attention

摘要: Metal surface defect detection has played a crucial role in the manufacturing of high-quality metal products. However, defect detection is challenging due to inter-class similarities and intra-class variances. Previous studies have primarily focused on detecting surface defects in individual metal categories, with limited exploration of models capable of detecting defects across multiple metal types. Aiming at the above issues, we proposed an improved model named DEFECT-YOLO, based on the YOLOv8s model. Concretely, a lightweight Effective Multi-scale Convolution Module (EMCM) is designed to enhance accuracy and reduce parameter count. Secondly, a multi-scale Atrous Spatial Pyramid Pooling with Efficient Multi-scale Attention (ASPPE) module is employed to enhance the ability of the model to capture local information. Subsequently, the multi-branch coordinate attention (MCA) mechanism is implemented in each prediction branch to further enhance the feature extraction capability for defects. Finally, the Normalized Wasserstein Distance (NWD) loss function is applied to jointly evaluate object localization loss with the complete intersection over union (CIoU) loss function, thereby improving the accuracy of the model for small-scale defects. Comprehensive experiments were conducted on a dataset containing various metal surface defects. The results indicate that the DEFECT-YOLO model achieves a 3.6% improvement in mean average precision (mAP@0.5) compared to the baseline model, with only a modest increase of 1.64 M parameters. Additionally, the model maintains real-time detection capabilities at 89 frames per second (FPS), which is essential for industrial applications. The code is available at https://github.com/zflyy/DEFECT-YOLO.git.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0920-8542  
eISSN:  
1573-0484  
入藏号:  
WOS:001337777100003  
Indexed Date:  
2024-10-26  
Institution Address:  
Xihua Univ, Sch Comp & Software Engn, Chengdu 610039, Peoples R China  
Institution Address:  
Chengdu Univ, Coll Chinese ASEAN Arts, Chengdu 610106, Peoples R China  
Institution:  
Institution:

### 1021. A Novel Semi-Supervised Graph-Guided Approach for Intelligent Health

摘要: As one of the most important operation and maintenance approaches, health state diagnosis technology plays a crucial role in ensuring the safety and reliability of mechanical equipment. The planar parallel manipulator, as a typical actuator, is widely employed in the field of precision manufacturing due to its advantages of high stiffness, large load support capability, and high precision. However, compared with common key functional components (such as bearings and gearboxes), planar parallel manipulators have more complicated operating mechanisms and failure behaviors. To satisfy the health state diagnosis demands of planar parallel manipulators in the scenario of insufficient label information, a novel intelligent health state diagnosis approach, termed semi-supervised graph-guided network with perception attention (SGN-PA), is developed for a 3-PRR (P and R represent prismatic and revolute pairs, respectively) planar parallel manipulator. Specifically, an improved multiorder graph perception module is constructed to extract multiscale feature information, and achieve feature fusion by combining perceptual attention mechanism, which enables the proposed SGN-PA model to have adaptation adjustment capabilities. Following that, local and nonlocal feature constraint strategies are employed with pseudo-label technology to reduce intraclass differences and maximize interclass differences, and then to fit the demands of health state diagnosis tasks. Eventually, based on the simulation and experimental scenarios of a 3-PRR planar parallel manipulator, the effectiveness and feasibility of the proposed SGN-PA model is extensively confirmed, and the diagnosis results show that it can significantly relax the constraints of label information while maintaining superior performances.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
被引频次合计:  
11  
ISSN:  
1083-4435  
eISSN:  
1941-014X  
入藏号:  
WOS:000788894900001  
Indexed Date:  
2022-05-08  
Institution Address:  
South China Univ Technol, Sch Mech & Automot Engn, Guangdong Key Lab Precis Equipment & Mfg Technol, Guangzhou 510640, Peoples R China  
Institution Address:  
Zhaoqing New Area Bur Adm Serv, Zhaoqing 526070, Peoples R China  
Institution:  
Institution:  
Zhaoqing New Area Bur Adm Serv

### 1022. MSF-YOLO: A multi-scale features fusion-based method for small object

摘要: Small object detection has been widely used in real-world applications, such as small object detection from the perspective of UAVs and industrial inspection to locate small defects visible on the surface of materials. The width of each layer of network structure is not enough to represent rich multi-scale information, which may result in the model being insensitive to small objects and low detection accuracy. To address the above issues, we propose an MSF-YOLO model on the basis of the YOLOv3 algorithm. First, the multi-scale features of image is fused. With respect to the original ResNet cell, the single convolutional scale is increased to four convolutional scales, and the features under each different perceptual field are fused to obtain rich hierarchical information from images. Second, the initial anchor box is optimized. Twice K-means clustering methods are invoked to optimize the size of the initial anchor box to improve the overlap of the anchor box, further improving the accuracy of the model. Finally, the convergence of model is accelerated. By introducing the weight parameters obtained from training on the COCO dataset, the training process of the model is optimized as well as the convergence of the model is accelerated. Experimental results on two public datasets show that MSF-YOLO outperforms YOLOv3 with an average accuracy of 98.67% and 97.51%, and performs very well in mAP and IoU metrics compared to state-of-the-art models. Finally, an industrial dataset is introduced for evaluation, and the results showed a 31.54% improvement over the original YOLOv3. In summary, the MSF-YOLO model proposed in this paper is adaptable to the small object detection task in many different scenarios.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001142490500005  
Indexed Date:  
2024-01-26  
Institution Address:  
Nanchang HangKong Univ, Dept Software, Nanchang, Jiangxi, Peoples R China  
Institution Address:  
Hongdu Aviat Ind Grp, Nanchang, Jiangxi, Peoples R China  
Institution Address:  
China Inst Water Resources & Hydropower Res, Beijing, Peoples R China  
Institution:  
Institution:  
Institution:

### 1023. Small target detection with remote sensing images based on an improved

摘要: IntroductionSmall target detection with remote sensing images is a challenging topic due to the small size of the targets, complex, and fuzzy backgrounds. MethodsIn this study, a new detection algorithm is proposed based on the YOLOv5s algorithm for small target detection. The data enhancement strategy based on the mosaic operation is applied to expand the remote image training sets so as to diversify the datasets. First, the lightweight and stable feature extraction module (LSM) and C3 modules are combined to form the feature extraction module, called as LCB module, to extract more features in the remote sensing images. Multi-scale feature fusion is realized based on the Res 2 unit, Dres 2, and Spatial Pyramid Pooling Small (SPPS) models, so that the receptive field can be increased to obtain more multi-scale global information based on Dres2 and retain the obtained feature information of the small targets accordingly. Furthermore, the input size and output size of the network are increased and set in different scales considering the relatively less target features in the remote images. Besides, the Efficient Intersection over Union (EIoU) loss is used as the loss function to increase the training convergence velocity of the model and improve the accurate regression of the model. Results and discussionThe DIOR-VAS and Visdrone2019 datasets are selected in the experiments, while the ablation and comparison experiments are performed with five popular target detection algorithms to verify the effectiveness of the proposed small target detection method.  
Web of Science 核心合集中的 "被引频次":  
8  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
9  
ISSN:  
1662-5218  
入藏号:  
WOS:000948531900001  
PubMed ID:  
36923945  
Indexed Date:  
2023-03-23  
Institution Address:  
Beihang Univ BUAA, Res Div 7, Beijing, Peoples R China  
Institution Address:  
Beihang Univ BUAA, Ctr Informat & Control, Sch Automation Sci & Elect Engn, Beijing, Peoples R China  
Institution Address:  
Shandong Agr & Engn Univ, Sch Informat Sci & Engn, Jinan, Peoples R China  
Institution:  
Institution:  
Institution:

### 1024. Attention enhanced machine instinctive vision with human-inspired

摘要: Salient object detection (SOD) enables machines to recognize and accurately segment visually prominent regions in images. Despite recent advancements, existing approaches often lack progressive fusion of low and high-level features, effective multi-scale feature handling, and precise boundary detection. Moreover, the robustness of these models under varied lighting conditions remains a concern. To overcome these challenges, we present Attention Enhanced Machine Instinctive Vision framework for SOD. The proposed framework leverages the strategy of Multi-stage Feature Refinement with Optimal Attentions-Driven Framework (MFRNet). The multi-level features are extracted from six stages of the EfficientNet-B7 backbone. This provides effective feature fusions of low and high-level details across various scales at the later stage of the framework. We introduce the Spatial-optimized Feature Attention (SOFA) module, which refines spatial features from three initial-stage feature maps. The extracted multi-scale features from the backbone are passed from the convolution feature transformation and spatial attention mechanisms to refine the low-level information. The SOFA module concatenates and upsamples these refined features, producing a comprehensive spatial representation of various levels. Moreover, the proposed Context-Aware Channel Refinement (CACR) module integrates dilated convolutions with optimized dilation rates followed by channel attention to capture multi- scale contextual information from the mature three layers. Furthermore, our progressive feature fusion strategy combines high-level semantic information and low-level spatial details through multiple residual connections, ensuring robust feature representation and effective gradient backpropagation. To enhance robustness, we train our network with augmented data featuring low and high brightness adjustments, improving its ability to handle diverse lighting conditions. Extensive experiments on four benchmark datasets - ECSSD, HKU-IS, DUTS, and PASCAL-S - validate the proposed framework's effectiveness, demonstrating superior performance compared to existing SOTA methods in the domain. Code, qualitative results, and trained weights will be available at the link: https://github.com/habib1402/MFRNet-SOD.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
7  
ISSN:  
0262-8856  
eISSN:  
1872-8138  
入藏号:  
WOS:001353398400001  
Indexed Date:  
2024-11-18  
Institution Address:  
Gachon Univ, Sch Comp, 1342 Seongnam daero, Seongnam 13120, South Korea  
Institution Address:  
Univ Technol Compiegne, Ctr Rech Royallieu, Lab Biomecan & Bioingn UMR 7338, Compiegne, France  
Institution:  
Institution:

### 1025. Fast fault detection method for photovoltaic arrays with adaptive deep

摘要: Photovoltaic (PV) arrays have output characteristics such as randomness and intermittency, and faults can seriously affect the safe operation of the power system. In order to improve the comprehensive performance of the PV array fault diagnosis model, a new intelligent online fault monitoring method for PV arrays is proposed in this paper. (1) a three-dimensional channel feature map based on I, V, and P features is constructed because the IV and P curves of the PV array have significantly different effects under different fault conditions. (2) The PV array fault diagnosis model based on a multi-source information fusion network (MIFNet) is proposed, and Channel Mixing Convolution (CMC) module, three-dimensional feature attention enhancement (TDFAE) module, and Channel normalized scaling (CNS) module are designed to improve the comprehensive performance of the model. (3) An adaptive nonlinear mutual sparrow search algorithm (ANMSSA) is proposed to optimize the hyperparameter configuration of the MIFNet network. The experimental results show that the average recognition accuracy, prediction accuracy, and sensitivity of the ANMSSA-MIFNet network proposed in this paper are 99.64%, 99.64%, and 99.71% respectively. When facing single-component faults and multi-component faults, the model has stronger diagnostic accuracy, robustness, anti-noise ability, and stability, and can efficiently diagnose different faults of PV arrays, providing the scientific basis and theoretical support for the operation of PV systems.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
16  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
16  
ISSN:  
0306-2619  
eISSN:  
1872-9118  
入藏号:  
WOS:001092910200001  
Indexed Date:  
2023-11-10  
Institution Address:  
Lanzhou Univ Technol, Coll Elect & Informat Engn, 36 Pengjiaping Rd, Lanzhou 730050, Peoples R China  
Institution Address:  
Lanzhou Univ Technol, Sch Comp & Commun, Lanzhou 730050, Peoples R China  
Institution Address:  
Lanzhou Univ Technol, Key Lab Gansu Adv Control Ind Proc, Lanzhou 730050, Peoples R China  
Institution Address:  
Lanzhou Univ Technol, Coll Energy & Power Engn, Lanzhou 730050, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 1026. Automatic Cobb angle measurement method based on vertebra segmentation

摘要: The accuracy of the Cobb measurement is essential for the diagnosis and treatment of scoliosis. Manual measurement is however influenced by the observer variability hence affecting progression evaluation. In this paper, we propose a fully automatic Cobb measurement method to address the accuracy issue of manual measurement. We improve the U-shaped network based on the multi-scale feature fusion to segment each vertebra. To enable multi-scale feature extraction, the convolution kernel of the U-shaped network is substituted by the Inception Block. To solve the problem of gradient disappearance caused by the widening of the network structure from the Inception Block, we propose using Res Block. CBAM (Convolutional Block Attention Module) can help the network judges the importance of the feature map to modify learning weight. Also, to further enhance the accuracy of feature extraction, we add the CBAM to the U-shaped network bottleneck. Finally, based on the segmented vertebrae, the efficient automatic Cobb angle measurement method is proposed to estimate the Cobb angle. In the experiments, 75 spinal X-ray images are tested. We compare the proposed U-Shaped network with the state-of-the-art methods including DeepLabV3 + , FCN8S, SegNet, U-Net, U-Net + + , BASNet, and U(2)Net for vertebra segmentation. Our results show that compared to these methods, the Dice coefficient is improved by 32.03%, 33.58%, 12.42%, 5.65%, 4.55%, 4.42%, and 3.27%, respectively. The CMAE of the calculated Cobb measurement is 2.45 degrees, which is lower than the average error of 5-7 degrees of manual measurement. The experimental results indicate that the improved U-shaped network improves the accuracy of vertebra segmentation. The proposed efficient automatic Cobb measurement method can be used in clinics to reduce observer variability.  
在 BIOSIS Citation Index 中的被引频次:  
6  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
22  
被引频次合计:  
23  
ISSN:  
0140-0118  
eISSN:  
1741-0444  
入藏号:  
WOS:000808422100001  
PubMed ID:  
35678952  
Indexed Date:  
2022-06-20  
Institution Address:  
Yunnan Univ, Sch Informat, East Outer Ring South Rd, Kunming 650504, Yunnan, Peoples R China  
Institution Address:  
Yunnan Prov Peoples Hosp, Peoples Hosp Yunnan 1, Orthoped Dept, 157 Jinbi Rd, Kunming 650034, Yunnan, Peoples R China  
Institution:  
Institution:  
Yunnan Prov Peoples Hosp

### 1027. Using Multi-Scale Convolutional Neural Network Based on Multi-Instance

摘要: Background: At present, radical total mesorectal excision after neoadjuvant chemoradiotherapy is crucial for locally advanced rectal cancer. Therefore, the use of histopathological images analysis technology to predict the efficacy of neoadjuvant chemoradiotherapy for rectal cancer is of great significance for the subsequent treatment of patients. Methods: In this study, we propose a new pathological images analysis method based on multi-instance learning to predict the efficacy of neoadjuvant chemoradiotherapy for rectal cancer. Specifically, we proposed a gated attention normalization mechanism based on the multilayer perceptron, which accelerates the convergence of stochastic gradient descent optimization and can speed up the training process. We also proposed a bilinear attention multi-scale feature fusion mechanism, which organically fuses the global features of the larger receptive fields and the detailed features of the smaller receptive fields and alleviates the problem of pathological images context information loss caused by block sampling. At the same time, we also designed a weighted loss function to alleviate the problem of imbalance between cancerous instances and normal instances. Results: We evaluated our method on a locally advanced rectal cancer dataset containing 150 whole slide images. In addition, to verify our method's generalization performance, we also tested on two publicly available datasets, Camelyon16 and MSKCC. The results show that the AUC values of our method on the Camelyon16 and MSKCC datasets reach 0.9337 and 0.9091, respectively. Conclusion: Our method has outstanding performance and advantages in predicting the efficacy of neoadjuvant chemoradiotherapy for rectal cancer. Clinical and Translational Impact Statement-This study aims to predict the efficacy of neoadjuvant chemoradiotherapy for rectal cancer to assist clinicians quickly diagnose and formulate personalized treatment plans for patients.  
Web of Science 核心合集中的 "被引频次":  
16  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
16  
ISSN:  
2168-2372  
入藏号:  
WOS:000770603200002  
PubMed ID:  
35317416  
Indexed Date:  
2022-03-30  
Institution Address:  
Yunnan Univ, Sch Software, Kunming 650106, Yunnan, Peoples R China  
Institution Address:  
Yunnan Univ, Sch Informat Sci & Engn, Kunming 650106, Yunnan, Peoples R China  
Institution Address:  
Yunnan Univ, Sch Software, Key Lab Software Engn Yunnan Prov, Kunming 650504, Yunnan, Peoples R China  
Institution Address:  
Kunming Med Univ, Yunnan Canc Hosp, Yunnan Canc Ctr, Affiliated Hosp 3, Kunming 650118, Yunnan, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 1028. Electrode Microscopic Image Segmentation Method by Fusing Multi-layer

摘要: To address the problems of blurred material edges,artifacts,and uneven grayscale in electrode microscopic images of NO\_x sensors,an electrode microscopic image semantic segmentation method that fuses multi-layer perceptual attention is proposed,in which U-Net is the base model. First,different scale output feature maps of the U-Net encoding layer with a 3\*3 convolution are used to reduce dimensionality. Furthermore,bilinear interpolation is used to unify feature scales to achieve multi-scale feature fusion,enhance feature information extraction,and compensate for feature loss from encoding downsampling. Second,by adding spatial pyramid pooling to extract multi-scale information and employing a 1\*1 convolution to reduce the calculation,a multi-layer perceptual attention module is proposed to capture the spatial position and channel dependence of the backbone feature map and the feature map with enhanced semantic information. Finally,a loss function with the ability to capture spatial similarity is proposed based on the similarity relationship of feature maps with different semantic information combined with cross-entropy loss. The key information is supervised during the training process to assist the backbone feature map to learn spatial position information and enhance the segmentation performance. The experimental results indicate that the Mean Pixel Accuracy(MPA) of the proposed method is 96.75%,the Mean Intersection over Union(MIoU) is 94.04%,Micro-F1 is 96.92%,FLOPs is 7.78\* 10~9,and the number of parameters contained in the network is 8.08\*10~6. Compared with models such as U-Net and SegNet,the proposed method can effectively address problems of edge blurring and material artifacts while increasing a little model complexity. Furthermore,it can capture spatial position and channel information,preserve detailed features of the image,and improve segmentation accuracy.  
摘要:  
针对氮氧传感器电极微观图像存在的物质边缘模糊、伪影、灰度不均等问题,将U-Net作为基础模型,提出融合多层感知注意力的电极微观图像语义分割方法。首先对U-Net编码层的不同尺度输出特征图使用3\*3卷积进行降维,利用双线性插值统一特征尺度,以实现多尺度特征融合,增强特征信息提取能力并补偿编码下采样中的特征损失;其次通过加入空间金字塔池化来提取多尺度信息并通过1\*1卷积减小计算量,同时提出多层感知注意力模块,以捕获主干特征图和增强语义信息特征图的空间位置与通道依赖关系;最后计算不同语义信息特征图的相似度关系,结合交叉熵损失提出具有捕获空间相似性能力的损失函数,在训练过程中对关键信息进行监督,辅助主干特征图学习空间位置信息,增强分割性能。实验结果表明,该方法的类别平均像素准确率为96.75%,平均交并比为94.04%,微观F1分数为96.92%,浮点运算次数为7.78\*10~9,网络所含参数量为8.08\*10~6。相对U-Net、SegNet等模型,该方法在提高少量模型复杂度的情况下,能有效改善边缘模糊及物质伪影问题,捕获空间位置与通道信息,保留图像细节特征,提高分割准确率。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1000-3428  
入藏号:  
CSCD:7639150  
Indexed Date:  
2024-04-04  
Institution Address:  
武汉科技大学计算机科学与技术学院;;智能信息处理与实时工业系统湖北省重点实验室, ;;智能信息处理与实时工业系统湖北省重点实验室, 武汉;;武汉, 湖北;;湖北 430065;;430065, 中国  
Institution Address:  
华中科技大学人工智能与自动化学院, 武汉, 湖北 430074, 中国  
Institution Address:  
College of Computer Science and Technology,Wuhan University of Science and Technology;;Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, ;;Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System, Wuhan;;Wuhan, Hubei;;Hubei 430065;;430065  
Institution Address:  
School of Artificial Intelligence and Automation,Huazhong University of Science and Technology, Wuhan, Hubei 430074, China  
Institution:  
武汉科技大学计算机科学与技术学院;;智能信息处理与实时工业系统湖北省重点实验室  
Institution:  
华中科技大学人工智能与自动化学院  
Institution:  
College of Computer Science and Technology,Wuhan University of Science and Technology;;Hubei Province Key Laboratory of Intelligent Information Processing and Real-time Industrial System  
Institution:  
School of Artificial Intelligence and Automation,Huazhong University of Science and Technology

### 1029. Lightweight infrared road scene detection model based on multiscale and

摘要: In view of occlusion and lack of texture details of infrared targets in road scenes, which leads to false detection and missed detection, a lightweight infrared road scene detection YOLO(You Only Look Once) model based on Multi-Scale and weighted Coordinate attention(MSC-YOLO) was proposed. YOLOv7-tiny was taken as the baseline model. Firstly, a multi-scale pyramid module PSA(Pyramid Split Attention) was introduced in different intermediate feature layers of the MobileNetV3, and a lightweight backbone extraction network MSM-Net(Multi-Scale Mobile Network) for multi-scale feature extraction was designed to solve the problem of feature pollution caused by the fixed-size convolution kernel, improving the fine-grained extraction ability of targets of different scales. Secondly, Weighted Coordinate Attention(WCA) mechanism was integrated into the feature fusion network, and the target position information obtained from the vertical and horizontal spatial directions of the intermediate feature map was superimposed to enhance the fusion ability of target features in different dimensions. Finally, the positioning loss function was replaced to Efficient Intersection over Union(EIoU) to calculate the length and width influencing factors of the predicted frame and the real frame separately, accelerating the convergence. The verification experiment was carried out on the Flir dataset. Compared with the YOLOv7-tiny model, the number of parameters is reduced by 67.3%, the number of floating-point operations is reduced by 54.6%, and the model size is reduced by 60.5% under the premise that mAP(IoU=0.5)(mean Average Precision(IoU=0.5)) is only reduced by 0.7 percentage points. The Frames Per Second(FPS) reaches 101 on the RTA 2080Ti, achieving a balance between detection performance and lightweight, and meets the real-time detection requirements of infrared road scenes.  
摘要:  
针对道路场景下红外目标遮挡、缺乏纹理细节而导致目标误检、漏检的问题,提出一种基于多尺度和加权坐标注意力的轻量化红外道路场景检测模型(MSC-YOLO)。以YOLOv7-tiny作为基线模型,首先,在MobileNetV3的不同中间特征层引入多尺度金字塔模块PSA(Pyramid Split Attention),设计一种多尺度特征提取的轻量化主干提取网络MSM-Net(Multi-Scale Mobile Network),解决固定大小卷积核造成的特征污染问题,提高对于不同尺度目标的细粒度提取能力;其次,在特征融合网络融入加权坐标注意力(WCA)机制,叠加从中间特征图垂直和水平空间方向上获取的目标位置信息,增强目标特征在不同维度上的融合能力;最后,替换定位损失函数为高效交并比(EIoU),分别计算预测框和真实框的长、宽影响因子,提高收敛速度。在Flir数据集上进行验证实验,与YOLOv7-tiny模型相比,在mAP(IoU=0.5)仅降低0.7个百分点的前提下,MSC-YOLO的参数量减少67.3%,浮点运算次数减少54.6%,模型大小减小60.5%,帧率在RTA 2080Ti上达到101,在检测性能和轻量化上达到平衡,满足红外道路场景的实时检测需求。  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1001-9081  
入藏号:  
CSCD:7740896  
Indexed Date:  
2024-08-23  
Institution Address:  
桂林理工大学信息科学与工程学院;;桂林理工大学, ;;广西嵌入式技术与智能系统重点实验室, 桂林;;桂林, 广西;;广西 541006;;541006, 中国  
Institution Address:  
桂林理工大学信息科学与工程学院, 桂林, 广西 541006, 中国  
Institution Address:  
桂林理工大学机械与控制工程学院, 桂林, 广西 541006, 中国  
Institution Address:  
College of Computer Science and Engineering, Guilin University of Technology;;Guilin University of Technology, ;;Guangxi Key Laboratory of Embedded Technology and Intelligent System, Guilin;;Guilin, Guangxi;;Guangxi 541006;;541006  
Institution Address:  
College of Computer Science and Engineering, Guilin University of Technology, Guilin, Guangxi 541006, China  
Institution Address:  
College of Mechanical and Control Engineering, Guilin University of Technology, Guilin, Guangxi 541006, China  
Institution:  
桂林理工大学信息科学与工程学院;;桂林理工大学  
Institution:  
桂林理工大学信息科学与工程学院  
Institution:  
桂林理工大学机械与控制工程学院  
Institution:  
College of Computer Science and Engineering, Guilin University of Technology;;Guilin University of Technology  
Institution:  
College of Computer Science and Engineering, Guilin University of Technology  
Institution:  
College of Mechanical and Control Engineering, Guilin University of Technology

### 1030. Intelligent Marine Survey: Lightweight Multi-Scale Attention Adaptive

摘要: Accurate and automatic underwater target recognition is a compelling challenge for autonomous underwater vehicles (AUVs) in intelligent marine surveys. This study proposed a seabed target correction model based on side-scan sonar (SSS) images and combined the navigation information of AUV to achieve pixel-level geocoding. Moreover, a lightweight multilevel attention adaptive segmentation framework (MA(2)Net) was proposed to achieve fine-grained recognition. It contains three new modules: 1) The lightweight attention network (LAN) is designed as the baseline to obtain dense feature maps and focus on interesting features based on a balanced attention mechanism. 2) the multi-scale feature pyramid (MASPP) was then constructed to capture the context of SSS images and extract rich semantic information at high levels. 3) Finally, the adaptive feature fusion module (AFF) effectively incorporates feature maps of MASPP and spatial information to improve the learned representations further. Extensive experiments are verified on six SSS categories and show the remarkable performance of the MA(2)Net compared with state-of-the-art methods. Furthermore, real sea trials were conducted by deploying MA(2)Net to the autonomous target recognition (ATR) system of AUV, which can achieve 29.7 fps and 81.23% MIoU for a (512 x 512) input on a single Nvidia Jetson Xavier.  
Note to Practitioners-This paper aims to provide a real-time semantic segmentation model for the autonomous target detection of AUV, which is suitable for the autonomous detection of underwater targets by underwater robots (ROV, AUV, ARV, et al). This paper proposes a lightweight, multi-scale attention-adaptive segmentation framework (MA(2)Net) incorporating pixel-level seabed targets rectification methods. The algorithm has high segmentation accuracy and fast operation speed. It can identify seabed targets in high-resolution sonar images online and realize precise positioning of small seabed targets, which is conducive to improving the intelligence level of marine survey unmanned equipment. This paper details the design of MA(2)Net and the hardware structure of the autonomous target recognition system (ATR). Plenty of simulation experiments and sea trials have proved the efficiency and practicability of the method for the autonomous detection of different seabed targets (sand waves, coral reefs, metal balls, threads, and artificial reefs). Future research will verify the generalization of the algorithm in more seabed targets.  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
1545-5955  
eISSN:  
1558-3783  
入藏号:  
WOS:001185941000001  
Indexed Date:  
2024-03-30  
Institution Address:  
Ocean Univ China, Sch Informat Sci & Engn, Qingdao 266000, Shandong, Peoples R China  
Institution:

### 1031. FSOU-Net: Feature supplement and optimization U-Net for 2D medical image

摘要: BACKGROUND: The results of medical image segmentation can provide reliable evidence for clinical diagnosis and treatment. The U-Net proposed previously has been widely used in the field of medical image segmentation. Its encoder extracts semantic features of different scales at different stages, but does not carry out special processing for semantic features of each scale.  
OBJECTIVE: To improve the feature expression ability and segmentation performance of U-Net, we proposed a feature supplement and optimization U-Net (FSOU-Net).  
METHODS: First, we put forward the view that semantic features of different scales should be treated differently. Based on this view, we classify the semantic features automatically extracted by encoders into two categories: shallow semantic features and deep semantic features. Then, we propose the shallow feature supplement module (SFSM), which obtains fine-grained semantic features through up-sampling to supplement the shallow semantic information. Finally, we propose the deep feature optimization module (DFOM), which uses the expansive convolution of different receptive fields to obtain multi-scale features and then performs multi-scale feature fusion to optimize the deep semantic information.  
RESULTS: The proposed model is experimented on three medical image segmentation public datasets, and the experimental results prove the correctness of the proposed idea. The segmentation performance of the model is higher than the advanced models for medical image segmentation. Compared with baseline network U-NET, the main index of Dice index is 0.75% higher on the RITE dataset, 2.3% higher on the Kvasir-SEG dataset, and 0.24% higher on the GlaS dataset.  
CONCLUSIONS: The proposed method can greatly improve the feature representation ability and segmentation performance of the model.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
3  
ISSN:  
0928-7329  
eISSN:  
1878-7401  
入藏号:  
WOS:000916321000016  
PubMed ID:  
35754242  
Indexed Date:  
2023-02-14  
Institution Address:  
Xinjiang Univ, Coll Software Engn, Urumqi, Xinjiang, Peoples R China  
Institution Address:  
Xinjiang Univ, Key Lab Software Engn Technol, Urumqi, Xinjiang, Peoples R China  
Institution Address:  
Xinjiang Univ, Coll Informat Sci & Engn, Urumqi, Xinjiang, Peoples R China  
Institution Address:  
Peoples Hosp Xinjiang Uygur Autonomous Reg, Urumqi, Xinjiang, Peoples R China  
Institution Address:  
China Xinjiang Key Lab Dermatol Res XJYS1707, Urumqi, Xinjiang, Peoples R China  
Institution Address:  
Sichuan Univ, Coll Comp Sci, Chengdu, Sichuan, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Peoples Hosp Xinjiang Uygur Autonomous Reg  
Institution:  
China Xinjiang Key Lab Dermatol Res XJYS1707  
Institution:

### 1032. LKR-DETR: small object detection in remote sensing images based on

摘要: Small object detection in remote sensing imagery remains a challenging problem in computer vision. To address the inherent limitations of aerial imagery, such as densely packed objects with insufficient detail and occlusions caused by complex backgrounds, this study proposes LKR-DETR, an innovative object detection in remote sensing imagery framework based on RT-DETR. We propose a lightweight and efficient feature extraction module with large kernel convolution, which expands the receptive field while reducing parameters and computational costs. Furthermore, we present a novel multi-scale feature fusion structure based on wavelet transform convolution that effectively utilizes low-frequency information from low-level feature maps. Additionally, we introduce a lightweight image restoration module utilizing large kernel convolutions, which effectively recovers previously undetected details of small objects. To improve bounding box regression accuracy, the original GIoU loss is replaced with a Focaler-DIoU loss function. Compared to the benchmark model RT-DETR, the LKR-DETR model achieves a 2.5% improvement in mAP0.5\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$mAP\_{0.5}$$\end{document} and a 2.0% improvement in mAP0.5:0.95\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$mAP\_{0.5:0.95}$$\end{document} on the VisDrone2019-DET dataset, a 1.7% and 4.4% improvement on the DOTAv1.5 dataset, and a 3.4% and 2.4% improvement on the HIT-UAV dataset, while also reducing the parameter count and model size. Relative to other cutting-edge models, LKR-DETR attains superior detection accuracy while maintaining relatively low computational complexity, establishing it as an efficient solution for small object detection in remote sensing imagery.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001400130100002  
Indexed Date:  
2025-01-24  
Institution Address:  
Dalian Minzu Univ, Coll Sci, Dalian 116000, Peoples R China  
Institution Address:  
Dalian Minzu Univ, Sch Preuniv, Dalian 116000, Peoples R China  
Institution:  
Institution:

### 1033. MFNet: Multi-scale feature enhancement networks for wheat head detection

摘要: Developing a high-precision wheat head detection algorithm is challenging due to the dense distribution and diverse sizes of wheat head in the field, as well as serious coverage from weeds. In this work, we propose multi-scale feature enhancement networks for wheat head detection and counting in complex scene (MFNet). Firstly, we introduce a deformable spatial attention mechanism (DSAM) and embed it in the backbone network to enhance the extraction of wheat head features while suppressing irrelevant features, effectively improving detection of wheat head in occluded environments. Secondly, we design a multi-scale receptive field feature fusion (MRFF) module in combination with an improved light-weight feature pyramid module to achieve more accurate detection of wheat head of different sizes and improve localization accuracy simultaneously. Additionally, the modified detection head with deformable convolution is able to adapt to different shapes of wheat head features and accurately predict the bounding boxes. Our method achieves 94.2% AP@50 at a speed of 30 FPS on the GWHD dataset, and to verify the generalization of the proposed method, we constructed a dense wheat head detection (DWHD) dataset with annotations, conducted experiments on the DWHD and SPIKE datasets and compared them with state-of-the-art algorithms. The experimental results show that the proposed method outperformed most of the existing methods, which further proved the superiority and robustness of the proposed method, and demonstrated that our model possesses excellent adaptability, enabling it to flexibly cope with scenarios of significant scale differences and severe occlusion in complex field environments for wheat head, and provide a technological reference for monitoring the wheat phenotype.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:001296869900001  
Indexed Date:  
2024-08-28  
Institution Address:  
Xinjiang Univ, Sch Software, Urumqi 830000, Peoples R China  
Institution Address:  
Key Lab Signal Detect & Proc Xinjiang Uygur Autono, Urumqi 830000, Peoples R China  
Institution Address:  
Xinjiang Univ, Key Lab Software Engn, Urumqi 830000, Peoples R China  
Institution Address:  
Xinjiang Univ, Coll Informat Sci & Engn, Urumqi 830000, Peoples R China  
Institution:  
Institution:  
Key Lab Signal Detect & Proc Xinjiang Uygur Autono  
Institution:  
Institution:

### 1034. Joint pyramid attention network for real-time semantic segmentation of

摘要: Semantic segmentation is an advanced research topic in computer vision and can be regarded as a fundamental technique for image understanding and analysis. However, most of the current semantic segmentation networks only focus on segmentation accuracy while ignoring the requirements for high processing speed and low computational complexity in mobile terminal fields such as autonomous driving systems, drone applications, and fingerprint recognition systems. Aiming at the problems that the current semantic segmentation task are facing, it is difficult to meet the actual industrial needs due to its high computational cost. We propose a joint pyramid attention network (JPANet) for real-time semantic segmentation. First, we propose a joint feature pyramid (JFP) module, which can combine multiple network stages with learning multi-scale feature representations with strong semantic information, hence improving pixel classification performance. Second, we built a spatial detail extraction (SDE) module to capture the shallow network multi-level local features and make up for the geometric information lost in the down-sampling stage. Finally, we design a bilateral feature fusion (BFF) module, which properly integrates spatial information and semantic information through a hybrid attention mechanism in spatial dimensions and channel dimensions, making full use of the correspondence between high-level features and low-level features. We conducted a series of experiments on two challenging urban road scene datasets (Cityscapes and CamVid) and achieved excellent results. Among them, the experimental results on the Cityscapes dataset show that for 512 x 1024 high-resolution images, our method achieves 71.62% Mean Intersection over Union (mIoU) with 109.9 frames per second (FPS) on a single 1080Ti GPU.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
40  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
4  
被引频次合计:  
44  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000648029400002  
Indexed Date:  
2021-06-01  
Institution Address:  
Chongqing Univ Posts & Telecommun, Key Lab Intelligent Anal & Decis Complex Syst, Chongqing 400065, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Multimedia Commun Res Lab, Chongqing 400065, Peoples R China  
Institution Address:  
Northeastern Univ, Cross Media Artificial Intelligence Lab, Shenyang 110000, Peoples R China  
Institution:  
Institution:  
Institution:

### 1035. An efficient feature aggregation network for small object detection in

摘要: Unmanned aerial vehicles (UAVs) possess high mobility and a wide field of view, leading to challenges such as a high proportion of small objects, significant variation in object size, object aggregation, and complex backgrounds in aerial images. Existing object detection methods often overlook the texture information in high-level features, which is crucial for detecting small objects in complex backgrounds. To improve the detection performance of small objects in complex scenes, we propose an efficient feature aggregation network (EFA-Net) based on YOLOv7. The backbone of the network seamlessly integrates a lightweight hybrid feature extraction module (LHFE), which replaces traditional convolutions with depthwise convolutions and employs a hybrid channel attention mechanism to capture local and global information concurrently. This design can effectively reduce the parameters without sacrificing detection accuracy and enhance the network's representative capacity. In the neck, we design an innovative adaptive multi-scale feature fusion module (AMSFM) that improves the model's adaptability to small objects and complex backgrounds by fusing multi-scale features with high-level semantic information and capturing the texture information in high-level features. Additionally, we incorporate a residual spatial pyramid pooling (RSPP) module to strengthen information fusion from various receptive fields and reduce the interference of complex backgrounds on small object detection. To further improve the model's robustness and generalization ability, we propose an enhanced complete intersection over union (ECIoU) loss function to balance the influence of large and small objects during training. Experimental results demonstrate the effectiveness of the proposed method, achieving mAP50\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$${mAP\_{50}}$$\end{document} scores of 51.6% and 48.5%, and mAP scores of 29.6% and 29.5% on the VisDrone 2019 and UAVDT datasets, respectively.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0920-8542  
eISSN:  
1573-0484  
入藏号:  
WOS:001434422700002  
Indexed Date:  
2025-03-05  
Institution Address:  
Zhengzhou Univ, Coll Comp Sci & Artificial Intelligence, Zhengzhou 450000, Henan, Peoples R China  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Comp Sci, Beijing 100876, Peoples R China  
Institution:  
Institution:

### 1036. MRT-YOLO: A Fine-Grained Feature-Based Method for Object Detection

摘要: Object detection is an essential component of autonomous driving, unmanned aerial vehicle (UAV) reconnaissance, and other domains. It equips drones and vehicles with the capability to perceive and comprehend their surrounding environment, making it a crucial technology for achieving safe and reliable autonomous driving as well as UAV spot reconnaissance. This paper proposes an end-to-end, high-precision, multi-scale real-time detection algorithm called MRT-YOLO based on YOLOv8. Firstly, in the feature downsampling process of the backbone network, we extend the channel depth to enhance the model's learning capability for fine-grained features and thereby improve its performance in retaining feature information. Secondly, we enhance the cross-stage partial layer version 2 (C2f) module in YOLOv8 by incorporating a channel self-attention mechanism within it, which optimizes performance through effective feature interaction and integration. Simultaneously, we also employ an improved bidirectional feature pyramid network (BiFPN) and introduce the proposed multi-scale feature learning (MFL) module to further enhance the model's feature extraction ability. In this study, we fuse the feature maps (C2, C3, C4, and C5) from the backbone network to generate a new feature map C6, thus increasing cross-connections between low-level and high-level features. Lastly, a multi-scale small object detection structure is designed to enhance recognition sensitivity toward densely distributed small objects. The proposed algorithm's effectiveness and superiority are demonstrated through experiments conducted on two datasets: VisDrone (UAV vision dataset) and BDD100K (automatic driving dataset).  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
2079-9292  
入藏号:  
WOS:001377823800001  
Indexed Date:  
2024-12-25  
Institution Address:  
Changsha Univ Sci & Technol, Sch Comp & Commun Engn, Changsha 410114, Peoples R China  
Institution Address:  
Hunan First Normal Univ, Sch Elect Informat, Changsha 410221, Peoples R China  
Institution:  
Institution:

### 1037. Progressive Multi-Scale Feature Cascade Fusion Color Constancy Algorithm

摘要: Color constancy is an important prerequisite for computer vision tasks such as object detection, three-dimensional object reconstruction, and automatic driving. In order to make full use of the feature information of different scales in the image to estimate the light source, a progressive multi-scale feature cascade fusion color constancy algorithm is proposed. The feature information in the image is extracted from different scales by three convolution network branches to fuse and get more abundant feature information. By cascading the shallow edge information and the deep fine-grained feature information in the image, the accuracy of the color constancy algorithm is improved. The progressive network structure improves the robustness of the algorithm for the light source estimation in extreme scenes by weighted cumulative angle error loss function. Experimental results on the reprocessed ColorChecker and NUS-8 datasets show that the proposed algorithm outperforms the current color constancy algorithm in terms of various evaluation indexes, and can be applied to other computer vision tasks requiring color constancy preprocessing.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
4  
ISSN:  
0253-2239  
入藏号:  
WOS:000828234200029  
Indexed Date:  
2022-07-29  
Institution Address:  
Beijing Inst Graph Commun, Sch Informat Engn, Beijing 102600, Peoples R China  
Institution:

### 1038. Focus on hierarchical features: Soft-weighted hierarchical features

摘要: Methods of multi-scale context are commonly used for semantic segmentation. Existing methods suffer from two flaws: (1) Underutilization of backbone-derived multi-scale feature information. (2) Mismatch between small objects and large-scale encodings. To solve the above issues, we propose a novel network to better represent high-level features, named as Soft-weighted Hierarchical Features Network (SWHF-Net) consisting of Semi-atrous Transform Feature Pyramid Module (ST-FPM) and Hierarchical Features Fusion Module (HF2M). Specifically, we propose a hierarchy-driven feature transformation function strat-egy to reconstruct the traditional feature pyramid module as ST-FPM. ST-FPM strengthens the properties of hierarchical features, which is beneficial for extracting the semantic representation of multi-scale objects. Simultaneously, HF2M focuses on the characteristics of features at different hierarchies, and adaptively calculates the attention map of multi-scale objects, greatly improving the efficiency. On Cityscapes, Pascal Context, and ADE20K, we achieve outstanding performance compared to the state-of-the-art methods with fewer computational costs.(c) 2022 Published by Elsevier B.V.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000883793500012  
Indexed Date:  
2022-11-26  
Institution Address:  
Xian Key Lab Big Data & Intelligent Vis, Xian 710071, Shaanxi, Peoples R China  
Institution Address:  
Xidian Univ, Sch Comp Sci & Technol, 2 Taibainan Rd, Xian 710071, Shaanxi, Peoples R China  
Institution Address:  
Xidian Univ, Guangzhou Inst technol, Guangzhou 510555, Guangdong, Peoples R China  
Institution Address:  
Guilin Univ Elect Technol, Guangxi Key Lab Trusted Software, Guilin 541004, Guangxi, Peoples R China  
Institution:  
Xian Key Lab Big Data & Intelligent Vis  
Institution:  
Institution:  
Institution:

### 1039. CSMViT: A Lightweight Transformer and CNN fusion Network for Lymph Node

摘要: To address the burdensome and time-consuming nature of manual diagnosis of pathological sections, this study proposes an automated pathological image detection system. This system can directly detect pathological images and accurately locate lesion tissues, providing a reference for pathological diagnosis. We propose an improved MobileViT model for feature extraction in the system, which we have named CSMViT. Considering the complexity and multi-scale characteristics of pathological images, we made three significant modifications to the MobileViT model. First, the original MV2 module was replaced with an improved Ghost module to reduce the model's parameter count, enhance detection accuracy, and accelerate inference speed. Second, we improved the backbone structure of the network to achieve multi-scale feature learning, which not only further reduces the parameter count but also allows for more effective capture of features at different scales. Lastly, we introduced a new CSA module that can simultaneously accept two feature maps of different sizes as input. Through internal attention mechanisms and feature fusion, this module achieves cross-scale feature learning. Experimental results indicate that the CSMViT model achieved accuracy, F1-score, and specificity of 99.42%, 99.4%, and 99.6%, respectively. Additionally, the detection accuracy of CSMViT for the entire pathological image is 84%, representing an 8% improvement over the original network. Notably, the FLOPs of CSMViT is 1.461G, which is a 72.19% reduction compared to the original network, significantly decreasing the model's complexity. These results thoroughly demonstrate the effectiveness and substantial value of CSMViT in pathological image detection.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001346084700001  
Indexed Date:  
2024-11-07  
Institution Address:  
Yantai Univ, Sch Phys & Elect Informat, Yantai 264005, Peoples R China  
Institution Address:  
Shandong First Med Univ & Shandong Acad Med Sci, Shandong Canc Hosp & Inst, Dept Radiat Oncol, Jinan 250117, Peoples R China  
Institution Address:  
Yantaishan Hosp, Dept Pathol, Yantai 264003, Peoples R China  
Institution:  
Institution:  
Institution:  
Yantaishan Hosp

### 1040. Lightweight YOLOv7 Algorithm for Multi-Object Recognition on Contrabands

摘要: With the strengthening of worldwide counter-terrorism initiatives, it is increasingly important to detect contrabands such as controlled knives and flammable materials hidden in clothes and bags. Terahertz (THz) imaging technology is widely used in the field of contraband detection due to its advantages of high imaging speed and strong penetration. However, the terahertz images are of poor qualities and lack texture details. Traditional target detection methods suffer from low detection speeds, misdetection, and omission of contraband. This work pre-processes the original dataset using a variety of image processing methods and validates the effect of these methods on the detection results of YOLOv7. Meanwhile, the lightweight and multi-object detection YOLOv7 (LWMD-YOLOv7) algorithm is proposed. Firstly, to meet the demand of real-time for multi-target detection, we propose the space-to-depth mobile (SPD\_Mobile) network as the lightweight feature extraction network. Secondly, the selective attention module large selective kernel (LSK) network is integrated into the output of the multi-scale feature map of the LWMD-YOLOv7 network, which enhances the effect of feature fusion and strengthens the network's attention to salient features. Finally, Distance Intersection over Union (DIOU) is used as the loss function to accelerate the convergence of the model and to have a better localisation effect for small targets. The experimental results show that the YOLOv7 algorithm achieves the best detection results on the terahertz image dataset after the non-local mean filtering process. The LWMD-YOLOv7 algorithm achieves a detection accuracy P of 98.5%, a recall R of 97.5%, and a detection speed of 112.4 FPS, which is 26.9 FPS higher than that of the YOLOv7 base network. The LWMD-YOLOv7 achieves a better balance between detection accuracy and detection speed. It provides a technological reference for the automated detection of contraband in terahertz images.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
eISSN:  
2076-3417  
入藏号:  
WOS:001171381800001  
Indexed Date:  
2024-03-08  
Institution Address:  
Henan Univ Technol, Key Lab Grain Informat Proc & Control, Minist Educ, Zhengzhou 450001, Peoples R China  
Institution Address:  
Henan Prov Key Lab Grain Photoelect Detect & Contr, Zhengzhou 450001, Peoples R China  
Institution Address:  
Henan Univ Technol, Coll Informat Sci & Engn, Zhengzhou 450001, Peoples R China  
Institution Address:  
Henan Univ Technol, Coll Artificial Intelligence & Big Data, Zhengzhou 450001, Peoples R China  
Institution:  
Institution:  
Henan Prov Key Lab Grain Photoelect Detect & Contr  
Institution:  
Institution:

### 1041. SNIS: A Signal Noise Separation-Based Network for Post-Processed Image

摘要: Image forgery detection has aroused widespread research interest in both academia and industry because of its potential security threats. Existing forgery detection methods achieve excellent tampered regions localization performance when forged images have not undergone post-processing, which can be detected by observing changes in the statistical features of images. However, forged images may be carefully post-processed to conceal forgery boundaries in a particular scenario. It becomes tough challenging to these methods. In this paper, we perform an analogous analysis between image forgery detection and blind signal separation, and formulate the post-processed image forgery detection problem into a signal noise separation problem. We also propose a signal noise separation-based (SNIS) network to solve the problem of detecting post-processed image forgery. Specifically, we first adopt the signal noise separation module to separate tampered region from the complex background region with post-processing noise, which weakens or even eliminates the negative impact of post-processing on forgery detection. Then, the multi-scale feature learning module uses a parallel atrous convolution architecture to learn high-level global features from multiple perspectives. Besides, a feature fusion module is utilized to enhance the discriminability of tampered regions and real regions by strengthening the boundary information. Finally, the prediction module is designed to predict the tampered region and classify the type of tampering operation. Extensive experiments show that the proposed SNIS is not only effective for forgery detection on forged images without post-processing, but also promising in robustness against multiple post-processing attacks. Furthermore, SNIS is robust in detecting forged images from unknown sources.  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
32  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
34  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:000941726100036  
Indexed Date:  
2023-03-27  
Institution Address:  
Hunan Univ, Coll Comp Sci & Elect Engn, Changsha 410082, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automat, Natl Lab Pattern Recognit, Beijing 100190, Peoples R China  
Institution Address:  
Fudan Univ, Sch Comp Sci, Shanghai 200433, Peoples R China  
Institution Address:  
Hunan Univ, Coll Elect & Informat Engn, Changsha 410082, Peoples R China  
Institution Address:  
Hunan Univ, Natl Engn Lab Robot Visual Percept & Control Techn, Changsha 410082, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 1042. Intelligent Retrieval of Radar Reflectivity Factor With Privacy

摘要: Meteorological radar data are essential for meteorological monitoring, forecasting, and research, and it plays a crucial role in observing and warning of extreme weather risks. However, meteorological radars have some limitations, such as uneven distribution and severe topographical influence. Meteorological remote sensing satellites can partially overcome these limitations by providing larger observational scope and high spatial and temporal resolution. Using data from meteorological remote sensing satellites to train radar reflectivity factor retrieval models can effectively compensate for the missing and poor quality of radar data. However, there are still some challenges, such as extracting the features of intense convective weather with unclear coverage from complex multichannel meteorological remote sensing satellite data and removing the interference caused by nonprecipitation clouds on retrieval models. Moreover, the privacy and security of remote sensing data transmission need to be ensured. In this article, we propose a novel method that combines the advanced encryption standard method to protect the transmission of remote sensing data, a multiscale feature fusion module to extract multiscale features from multichannel meteorological remote sensing satellite data, and an attention technique to reduce the interference of nonprecipitation clouds on retrieval models. We conduct comparison experiments with multiple indicators to demonstrate that our method has certain advantages in retrieving radar reflectivity values of different sizes. Our method achieves 0.63, 0.36, 0.49, 0.55, and 0.99 on probability of detection, false alarm ratio, critical success index, Heidke skill score, and accuracy scores, respectively.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001041983500008  
Indexed Date:  
2023-08-17  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Software, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Jiangsu Collaborat Innovat Ctr Atmospher Environm, Nanjing 210044, Peoples R China  
Institution Address:  
Hankuk Univ Foreign Studies, Dept Comp & Elect Syst Engn, Yongin 17579, South Korea  
Institution Address:  
Nanjing Meteorol Bur, Nanjing 210044, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Nanjing Meteorol Bur

### 1043. SmallNet: A Small Defects Detection Network for Magnetic Chips Based on

摘要: Accurate detection of surface defects on magnetic chips is a necessary and difficult task, especially for small defects, which lack discriminative and robust features due to their small size and weak characteristics. At present, the low accuracy of small defect detection seriously restricts the development of automated visual inspection and needs to be solved urgently. Thus, we propose a novel small defects detection network based on context-weighted aggregation and feature multiscale loop fusion. First, a context-weighted aggregation module (CAM) that enriches feature representations by combining context and attention mechanisms is proposed. We believe that any further improvements will be futile if robust feature representations cannot be obtained. Second, inspired by the mechanism of looking and thinking twice, we aggregate left-right feedback connections into feature pyramids and creatively propose a loop-shaped feature pyramid network (Loop-FPN), enabling multiscale features to be fused up and down and connected left and right. This loop-shaped structure makes the connection of each layer more direct and allows the features at each scale to be fully integrated, which improves the utilization of multiscale features and facilitates the detection of small defects. Finally, we apply the proposed network to practical detection and the results show that our network achieves 97.57% precision, 91.91% recall, and 98.39% AP, which are 0.13%, 1.43%, and 0.99% higher than the current best-performing comparative methods, respectively. Note to Practitioners-Current vision inspection technology has low accuracy and poor stability in small defect detection, which cannot meet the industrial inspection requirements. Small defect detection has become the biggest challenge in the field of visual inspection and has seriously restricted its development. Our proposed network can solve this problem well by mining small defect context, fusing multiscale features and utilizing attention mechanisms, and has been successfully applied in magnetic chip production line. Furthermore, we developed an image acquisition system that can capture defects on all surfaces with high accuracy and without dead space, allowing our network to not only detect small defects on magnetic chips of varying sizes and irregular shapes, but also to adapt to changes in lighting conditions, backgrounds, and viewpoints. Our work provides an effective, reliable, and convenient quality control solution for magnetic chip production.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-5955  
eISSN:  
1558-3783  
入藏号:  
WOS:001381483800001  
Indexed Date:  
2024-12-27  
Institution Address:  
Tongji Univ, Dept Comp Sci & Technol, Shanghai 201804, Peoples R China  
Institution:

### 1044. A deep-learning model with improved capsule networks and LSTM filters

摘要: The deep-learning networks for bearing fault diagnosis may encounter ambient noise interference. The inherently feedforward serial structure lacks the suppression of the ambient noise involved in bearing data. Moreover, the accuracy of the diagnosis model relies on training a large amount of labeled data, thus resulting in a significant amount of time consumption. To combat these two challenges, an improved capsule network diagnosis model with a long short-term memory filter (LF-iCapsNet) is proposed. First, the LSTM network is used to filter out the noise interference in the time domain through a nonlinear moving-average mechanism. At the same time, a subsequent feature-extraction convolution with large kernels is employed to specifically suppress the noise in the frequency domain, where a dilated convolution is adopted to achieve multi-scale feature extraction. Second, to obtain the distance-dependent relations of low-level features, an inner dependence operator is introduced into the primary capsule. Thus, a richer and more complete feature description of the bearing fault is guaranteed. And so, the derived digital capsule gives a geometrical constraint representation among the pixels in the feature maps and a rather high-speed training process because of its one-stage detection mode. Finally, the proposed LF-iCapsNet model is validated on the dataset from Case Western Reserve University (CWRU). The experimental results show that the diagnosis model provides a considerable improvement in the classification accuracy of bearing faults under the noise condition.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
8  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:000846669600001  
Indexed Date:  
2022-09-03  
Institution Address:  
Henan Polytech Univ, Sch Elect Engn & Automat, Henan Key Lab Intelligent Detect & Control Coal M, Jiaozuo 454003, Henan, Peoples R China  
Institution:

### 1045. ADS-YOLO: A Multi-Scale Feature Extraction Remote Sensing Image Object

摘要: Object detection in remote sensing images is of significant research value in fields such as environmental monitoring and urban planning. However, the large variation in object sizes, along with challenges such as small and densely packed objects, makes this task particularly challenging. To address these issues, we propose an algorithm for multi-scale feature extraction in remote sensing image detection using dilated residuals (ADS-YOLO). Firstly, to address the challenges of scale variation and small target size, the Dilation-wise Residual (DWR) design is employed to form the C2f\_DWR module, which restructures the bottleneck structure within the C2f segment to facilitate the extraction and fusion of multi-scale contextual information, thus reducing the difficulty associated with target scale variation. Secondly, inspired by the Adown subsampling convolution module from YOLOv9, we use it to replace the convolutions in the Backbone, enabling the model to capture finer image details at higher levels, while maintaining accuracy and reducing computational load. Lastly, to address the issue of dense targets, we design the Soft-NMS-ShapeIoU module to improve the consistency of bounding boxes and target shapes, while also suppressing adjacent boxes. Experimental results demonstrate that, on the publicly available remote sensing image datasets DIOR, RSOD, and NWPU VHR-10, the proposed ADS-YOLO model outperforms other state-of-the-art methods by a significant margin.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001422034400041  
Indexed Date:  
2025-02-21  
Institution Address:  
China West Normal Univ, Sch Elect Informat Engn, Nanchong 637009, Peoples R China  
Institution:

### 1046. Multilevel feature aggregation and enhancement network for remote

摘要: . Remote sensing change detection refers to the process of identifying and extracting changes in objects within the same geographical region over multiple periods. With the increasing spatial resolution of remote sensing images, the detection of minor changes has become a challenging task. We introduce a multilevel feature aggregation and enhancement network to tackle this issue. Specifically, we propose a multilevel feature aggregation module to aggregate the distinct features extracted from each image, which strengthens the feature representation capability. Subsequently, a difference parallel mapping module is designed to perceive information at different scales by refining the fused features. In addition, our guided change enhancement module captures local and long-range dependencies in multilevel features, improving the network's accuracy in identifying changing regions. Based on a basic shared weight Siamese backbone without complex structures, our model outperforms other state-of-the-art methods on three datasets in terms of both efficiency and effectiveness.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
4  
eISSN:  
1931-3195  
入藏号:  
WOS:001271750200051  
Indexed Date:  
2024-08-04  
Institution Address:  
Shandong Univ, Sch Software, Jinan, Peoples R China  
Institution:

### 1047. Urban Street Scene Instance Segmentation: An Integrated Hybrid Network

摘要: There are two standard methods in instance segmentation: top -down and bottom -up. The top -down approach performs object detection to generate candidate proposals and then performs pixel -level segmentation for each proposal. It is accurate and flexible, capable of handling objects of different sizes and shapes. However, it is computationally complex and relies on object detection accuracy. The bottom -up approach first performs pixel -level clustering or segmentation and then combines candidate instances to obtain the final segmentation result. It can handle overlapping cases and has lower computational complexity, but it may need to localize accurately, and segment instances, and the segmentation granularity is coarser. In this paper, the Urban Street Scene Instance Segmentation (UISNet) algorithm is proposed. Firstly, the feature extraction network is the foundation of UISNet, which uses EfficientNet as the backbone network. Secondly, MPAFPN is the feature pyramid network part of UISNet, used for multi -scale feature fusion. By using EfficientNet and MPAFPN as the backbone network and bottleneck layers, the accuracy of UISNet is improved by 4% compared to ResNet and FPN. In the inference phase, this paper introduces an innovative dual -branch design that combines top -down and bottom -up strategies. One branch is the bounding box aggregation branch, which generates high dimensional information such as the shape and orientation of bounding boxes based on the FCOS Head. The other branch is the mask decoding branch, which creates mask prediction results. These two branches are fused using the Mask FCN Header to obtain the final instance segmentation result. With this dual -branch design, the model can effectively utilize the information from both top -down and bottom -up approaches, thereby improving the accuracy and robustness of instance segmentation. Through experimental comparisons, the proposed network model in this paper achieves the best performance in terms of accuracy compared to other instance segmentation networks, with an accuracy of 36.28%. Moreover, the proposed model performs better in urban street scenes, enhancing object detection and segmentation and offering more reliable and efficient solutions for applications such as autonomous driving and intelligent transportation.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1816-093X  
eISSN:  
1816-0948  
入藏号:  
WOS:001252651400017  
Indexed Date:  
2024-07-18  
Institution Address:  
Univ Sci & Technol Liaoning, Sch Comp Sci & Software Engn, Anshan, Peoples R China  
Institution:

### 1048. Chinese named entity recognition for agricultural diseases based on

摘要: Named entity recognition is a crucial step in information extraction for agricultural diseases. However, most existing works only utilize word embedding models to generate contextual semantic features, which are limited by textual contextual dependencies and suffer from the problem of missing contextual semantic features. There is a semantic correlation between the visual information in agricultural disease images and the target entities in agricultural disease texts, but this visual information has not been fully utilized among the existing works. To solve the aforementioned issues, we propose a Chinese named entity recognition model for agricultural diseases based on entity-related visual prompts injection. First, we introduce an improved object detector with Shape-IoU to obtain local images related to the target entities. Second, we propose a visual feature extraction structure based on multi-scale feature fusion to extract visual features from the local images and encode them as visual prompts. Third, we propose a novel visual-guided attention mechanism to achieve multimodal information fusion within the pretrained language model. Furthermore, the first multimodal named entity recognition (MNER) dataset for agricultural diseases named Disease7000 was also collected and annotated, which contains 7,000 image-text pairs, 4 entity types, and 31,656 samples. Experimental results show that our proposed model achieved the best F1 scores of 89.17% and 75.06% on Disease7000 and the publicly available Twitter2015 dataset, respectively, demonstrating the effectiveness and generalizability of the model. Additionally, the ablation study validate the effectiveness of each module in many aspects.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:001372264100001  
Indexed Date:  
2024-12-13  
Institution Address:  
Hebei Agr Univ, Sch Informat Sci & Technol, Baoding 071001, Peoples R China  
Institution Address:  
Natl Engn Res Ctr Informat Technol Agr, Beijing 100097, Peoples R China  
Institution Address:  
Beijing Res Ctr Informat Technol Agr, Beijing 100097, Peoples R China  
Institution Address:  
Hebei Key Lab Agr Big Data, Baoding 071001, Peoples R China  
Institution Address:  
High resolut Earth Observat Syst Data & Applicat T, Baoding 071001, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Hebei Key Lab Agr Big Data  
Institution:  
High resolut Earth Observat Syst Data & Applicat T

### 1049. Evolving traffic sign detection via multi-scale feature enhancement,

摘要: Accurate and fast detection of traffic signs is critical to the navigation and safety in autonomous driving, especially for scenarios characterized by a diverse range of sign sizes and detection distances. Existing approaches, incorporating attention modules or modifying detection heads, frequently lead to high rates of false and missed detections due to the increased sampling depth. To address these deficiencies, we introduce a novel object detection framework, ERF-YOLO, where the Adaptive Sequence Fusion coordinates diversified features to amplify its fusion accuracy of localization and semantic information, the Self-calibration Feature Enhancement dynamically adjusts the original and fused features to enhance their feature expression, and the Wavelet-based Feature Reconstruction adopts a 2D feature decomposition and reconstruction to extract dominant features while weaken high-frequency noises. Comparative studies on the TT100K and CCTSDB 2021 datasets demonstrate a superior performance of ERF-YOLO, whose mAP achieve 84.2% and 84%, with significant improvements over its benchmarks at 5.5% and 1.4%, respectively.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1051-2004  
eISSN:  
1095-4333  
入藏号:  
WOS:001420287400001  
Indexed Date:  
2025-02-20  
Institution Address:  
Qingdao Univ, Sch Automat, Qingdao, Peoples R China  
Institution Address:  
Qingdao Univ, Shandong Key Lab Ind Control Technol, Qingdao, Peoples R China  
Institution Address:  
Beijing Normal Univ, Inst AI & Future Networks, Zhuhai, Peoples R China  
Institution:  
Institution:  
Institution:

### 1050. Powerful embedding networks for few-shot image classification

摘要: Few-shot image classification commits to recognizing new concepts from limited annotated samples. Our insight is to obtain a sufficiently powerful embedding network (PEN) to solve few-shot classification tasks. We propose a method to tackle the few-shot classification tasks, namely PENs for few-shot image classification. The key core of PEN is gaining a well-trained embedding network that is capable of extracting strong discriminating representations to represent an image by utilizing two strategies. One strategy is that the multi-scale feature maps are fused instead of only utilizing the final top-level feature maps. We consider that low-level features also play an important role instead of only utilizing top-level representations. Another significant strategy is knowledge distillation (KD). The characteristics of KD can help us get better performance of an embedding network to extract features. Finally, a distance function is employed to classify unlabeled samples. Comprehensive experiments are conducted on few-shot benchmarks. Our method achieves promising performances. The results demonstrate that KD and future fusion are beneficial to gain an expected embedding network for few-shot classification tasks. (C) 2021 SPIE and IS&T  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:000737383500009  
Indexed Date:  
2022-01-09  
Institution Address:  
Wuhan Univ, Elect Informat Sch, Wuhan, Peoples R China  
Institution:

### 1051. Real-time detection of underwater river crab based on multi-scale

摘要: To achieve rapid and accurate detection of pond cultured river crab, an underwater river crab target detection method based on multi-scale pyramid fusion image enhancement and the MobileCenterNet model is proposed in this paper. Firstly, considering the characteristics of underwater crab image, such as blurred and uneven brightness, multi-scale pyramid fusion is proposed to enhance underwater crab image based on Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance contrast and Underwater Dark Channel Prior (UDCP) for fog removal. Secondly, a crab target detection method based on the MobileCenterNet model is presented. The improved MobileNetv2 backbone network with coordinate attention module is applied to extract crab features, which not only achieves lightweight but also focuses the model's attention on crab-related features. Then, the Feature Fusion Module (FFM) is designed to extract multi-scale feature map information, and Atrous Spatial Pyramid Pooling (ASPP) is added to fuse context information from different receptive fields. The experimental results show that the average precision (AP) and F1 values of MobileCenterNet are 97.86 % and 97.94 %, the model size is only 24.46 M, and the detection speed is reaching 48.18 frames/s. Compared with the baseline model ResNet18-CenterNet, the storage memory required for model training is reduced by 81 %, and the AP on the crab dataset is increased by 3.2 %. The experimental results show that the proposed method can achieve real-time and accurate detection of underwater river crabs, and provide effective guidance for real-time monitoring and scientific feeding in crab breeding.  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
9  
Web of Science 核心合集中的 "被引频次":  
34  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
35  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:000900074100002  
Indexed Date:  
2023-01-08  
Institution Address:  
Jiangsu Univ, Sch Elect & Informat Engn, Zhenjiang 212013, Peoples R China  
Institution:

### 1052. Interactive multi-scale feature representation enhancement for small

摘要: In the field of detection, there is a wide gap between the performance of small objects and that of medium, large objects. Some studies show that this gap is due to the contradiction between the classification-based backbone and localization. Although the reduction in the feature map size is beneficial for the extraction of abstract features, it will cause the loss of detailed features in the localization as traversing the backbone. Therefore, an interactive multi-scale feature representation enhancement strategy is proposed. This strategy includes two modules: first a multi-scale auxiliary enhancement network is proposed for feature interaction under multiple inputs. We scale the input to multiple scales corresponding to the prediction layers, and only passes through the lightweight extraction module to extract more detailed features for enhancing the original futures. Moreover, an adaptive interaction module is designed to aggregate the features of adjacent layers. This approach provides flexibility in achieving the improvement of small objects detection ability without changing the original network structure. Comprehensive experimental results based on PASCAL VOC and MS COCO datasets show the effectiveness of the proposed method. ? 2021 Elsevier B.V. All rights reserved.  
Web of Science 核心合集中的 "被引频次":  
14  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
15  
ISSN:  
0262-8856  
eISSN:  
1872-8138  
入藏号:  
WOS:000639853600008  
Indexed Date:  
2021-05-13  
Institution Address:  
Jiangnan Univ, Key Lab Adv Proc Control Light Ind Minist Educ, Wuxi 214122, Jiangsu, Peoples R China  
Institution:

### 1053. Cross-scale information enhancement for object detection

摘要: Object detection usually adopts multi-scale fusion to enrich the information of the object, and the Feature Pyramid Network (FPN) is a common method for multi-scale fusion. However, traditional fusion methods such as FPN cause information loss when fusing high-level feature maps with low-level feature maps. To solve these problems, we propose a simple but effective cross-scale fusion method that fully uses the information of multi-scale feature maps. In addition, to better utilize the multi-scale contextual information, we designed the Selective Information Enhancement (SIE) module. The SIE dynamically selects information at more important scales for objects of different size and fuse the selected information with feature maps for information enhancement. Apply our method to Single Shot Multibox Detector (SSD) and propose a Cross-Scale Information Enhancement Single Shot Multibox Detector (CESSD). The CESSD improves the object detection capability of SSD models by fusing multi-scale features and selectively enhancing feature map information. To evaluate the effectiveness of the model, we validated it on the Pascal VOC2007 test set for 300 x 300 inputs, and the mean Average Precision (mAP) of CESSD reached 79.8%.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001173851000008  
Indexed Date:  
2024-03-22  
Institution Address:  
Shenyang Univ Chem Technol, Equipment Reliabil Inst, Shenyang 110142, Peoples R China  
Institution Address:  
Shenyang Univ Chem Technol, Mech & Power Engn Coll, Shenyang 110142, Peoples R China  
Institution:  
Institution:

### 1054. Warp-based edge feature reinforcement network for medical image

摘要: Background Rapid and accurate segmentation of medical images can provide important guidance in the early stages of life-threatening diseases. Purpose However, fuzzy edges and high similarity with the background in images usually cause undersegmentation or oversegmentation. To solve these problems. Methods We propose a novel edge features-reinforcement (EFR) module that uses relative frequency changes before and after warping images to extract edge information. Then, the EFR module leverages deep features to guide shallow features to produce a band-shaped edge attention map for reinforcing the edge region of all channels. We also propose a multiscale context exploration (MCE) module to fuse multiscale features and to extract channel and spatial correlations, which allows a model to focus on the parts that contribute most to the final segmentation. We construct EFR-Net by embedding EFR and MCE modules on the encoder-decoder architecture. Results We verify EFR-Net's performance with four medical datasets: retinal vessel segmentation dataset DRIVE, endoscopic polyp segmentation dataset CVC-ClinicDB, dermoscopic image dataset ISIC2018, and aortic true lumen dataset Aorta-computed tomography (CT). The proposed model achieves Dice similarity coefficients (DSCs) of 81.61%, 92.87%, 89.87%, and 96.98% on DRIVE, CVC-ClinicDB, ISIC2018, and Aorta-CT, respectively, which are better than those of current mainstream methods. In particular, the DSC of polyp segmentation increased by 3.87%. Conclusion Through quantitative and qualitative research, our method is determined to surpass current mainstream segmentation methods, and EFR modules can effectively improve the edge prediction effect of color images and CT images. The proposed modules are easily embedded in other encoder-decoder architectures, which has the potential to be applied and expanded.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0094-2405  
eISSN:  
2473-4209  
入藏号:  
WOS:000832979800001  
PubMed ID:  
35870115  
Indexed Date:  
2022-08-06  
Institution Address:  
Xinjiang Univ, Sch Informat Sci & Engn, Urumqi, Peoples R China  
Institution Address:  
Xinjiang Univ, Peoples Hosp Xinjiang Uygur Autonomous Reg, Urumqi, Peoples R China  
Institution Address:  
Xinjiang Univ, Coll Software, Urumqi 830000, Peoples R China  
Institution Address:  
Xin Jiang Univ, Coll Software, Key Lab Software Engn Technol, Urumqi, Peoples R China  
Institution Address:  
Xinjiang Univ, Coll Informat Sci & Engn, Signal & Signal Proc Lab, Urumqi, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 1055. MCFSA-Net: A multi-scale channel fusion and spatial activation network

摘要: As the only vascular tissue that can be directly viewed in vivo, retinal vessels are medically important in assisting the diagnosis of ocular and cardiovascular diseases. They generally appear as different morphologies and uneven thickness in fundus images. Therefore, the single-scale segmentation method may fail to capture abundant morphological features, suffering from the deterioration in vessel segmentation, especially for tiny vessels. To alleviate this issue, we propose a multi-scale channel fusion and spatial activation network (MCFSA-Net) for retinal vessel segmentation with emphasis on tiny ones. Specifically, the Hybrid Convolution-DropBlock (HC-Drop) is first used to extract deep features of vessels and construct multi-scale feature maps by progressive down-sampling. Then, the Channel Cooperative Attention Fusion (CCAF) module is designed to handle different morphological vessels in a multi-scale manner. Finally, the Global Spatial Activation (GSA) module is introduced to aggregate global feature information for improving the attention on tiny vessels in the spatial domain and realizing effective segmentation for them. Experiments are carried out on three datasets including DRIVE, CHASE\_DB1, and STARE. Our retinal vessel segmentation method achieves Accuracy of 96.95%, 97.57%, and 97.83%, and F1 score of 82.67%, 81.82%, and 82.95% in the above datasets, respectively. Qualitative and quantitative analysis show that the proposed method outperforms current advanced vessel segmentation methods, especially for tiny vessels.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1864-063X  
eISSN:  
1864-0648  
入藏号:  
WOS:000892949200001  
PubMed ID:  
36413066  
Indexed Date:  
2022-12-18  
Institution Address:  
Shandong Univ Sci & Technol, Coll Elect & Informat Engn, Qingdao 266590, Peoples R China  
Institution Address:  
Minjiang Univ, Coll Comp & Control Engn, Fujian Prov Key Lab Informat Proc & Intelligent Co, Fuzhou, Peoples R China  
Institution Address:  
Zhengzhou Univ, Sch Comp & Artificial Intelligence, Zhengzhou, Peoples R China  
Institution Address:  
Minjiang Univ, Fuzhou Digital Healthcare Ind Technol Innovat Ctr, Fuzhou 266590, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 1056. DRI-UNet: dense residual-inception UNet for nuclei identification in

摘要: Nuclei segmentation has great significance in biomedical applications as the preliminary step for disease diagnosis and treatment analysis. In this study, we propose a model for automated nuclei identification of varying cell shapes and types from microscopy images. Identifying nuclei helps to understand the underlying mechanism of various diseases in their early stages and provides solutions to enable faster cures. The foremost aim of the study is to develop a lightweight model, capable of segmenting varied shapes and sizes. The proposed architecture exploits multi-scale low-level features following dense high-level feature extraction with multi-feature fusion and special skip connections resulting in enhanced learning capability. The multi-scale feature extractor module extracts low-level information which is further processed using attention-based dense connections to extract semantically meaningful information. The special short-skip residual connections replacing long-skip connections reduced the semantic gap between encoder-decoder features. Moreover, the context encoder module extracts higher-level contextual information of different receptive fields using dilated convolutions making the model robust to different shapes and sizes. The higher-level feature maps propagate upward the decoder connections following the shared attention mechanism of an encoder to decoder features to reconstruct a better segmentation map. Moreover, the evaluation scheme following the proposed test-time augmentation operations improved the mean segmentation performance. The experiments on KDSB18, Synthetic cells, Triple-negative breast cancer (TNBC), MoNuSeg, CryoNuSeg, and BUS datasets demonstrate the suitability of the model for the nuclei segmentation tasks. The DRI-UNet model holds good segmentation performance outperforming baseline architecture by 8.12%, 4.71%, 10.19%, 2.46%, 3.14%, 8.91%, and 9.32% on KDSB18, synthetic cells, TNBC, MoNuSeg, CryoNuSeg, CVC-ClinicDB, and BUS datasets, respectively. We further conducted generalization tests of the proposed model for cross-dataset validation, and two independent MIS datasets confirm model effectiveness for nuclei cell and biomedical image segmentation.  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
被引频次合计:  
4  
ISSN:  
0941-0643  
eISSN:  
1433-3058  
入藏号:  
WOS:001014497100004  
Indexed Date:  
2023-07-27  
Institution Address:  
Banaras Hindu Univ, Inst Sci, Dept Comp Sci, Varanasi 221005, India  
Institution:

### 1057. Deep saliency detection-based pedestrian detection with multispectral

摘要: In recent years, there has been increased interest in multispectral pedestrian detection using visible and infrared image pairs. This is due to the complementary visual information provided by these modalities, which enhances the robustness and reliability of pedestrian detection systems. However, current research in multispectral pedestrian detection faces the challenge of effectively integrating different modalities to reduce miss rates in the system. This article presents an improved method for multispectral pedestrian detection. The method utilises a saliency detection technique to modify the infrared image and obtain an infrared-enhanced map with clear pedestrian features. Subsequently, a multiscale image features fusion network is designed to efficiently fuse visible and IR-enhanced maps. Finally, the fusion network is supervised by three loss functions for illumination perception, light intensity, and texture information in conjunction with the light perception sub-network. The experimental results demonstrate that the proposed method improves the logarithmic mean miss rate for the three main subgroups (all day, day and night) to 3.12%, 3.06%, and 4.13% respectively, at "reasonable" settings. This is an improvement over the traditional method, which achieved rates of 3.11%, 2.77%, and 2.56% respectively, thus demonstrating the effectiveness of the proposed method.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2296-424X  
入藏号:  
WOS:001154628700001  
Indexed Date:  
2024-02-09  
Institution Address:  
Xian Univ Sci & Technol, Coll Commun & Informat Engn, Xian, Peoples R China  
Institution Address:  
Xian Key Lab Heterogeneous Network Convergence Com, Xian, Peoples R China  
Institution Address:  
Shaanxi Cuijiagou Energy Co Ltd, Safety Supervis Dept, Tongchuan, Peoples R China  
Institution:  
Institution:  
Xian Key Lab Heterogeneous Network Convergence Com  
Institution:  
Shaanxi Cuijiagou Energy Co Ltd

### 1058. Vehicle Detection Based on Improved SSD Algorithm

摘要: The Single Shot Multibox Detector(SSD)algorithm uses multi-scale feature maps for classification and position regression.This algorithm displays better performance than the YOLO algorithm in detecting small targets,but frequently misses targets in vehicle detection.To address the problem,this paper proposes an improved SSD algorithm.In order to extract more vehicle feature information,an improved Inception module is given to replace the Conv8, Conv9 and Conv10 layers in the SSD network.At the same time,a network is designed to balance and fuse multi-scale features.The network can fuse the location information of shallow features and the semantic information of deep features,so the accuracy of small target detection can be increased.To further improve model performance,SENet is introduced into feature extraction layers to recalibrate the importance of different feature channels.The experimental results show that the improved SSD algorithm displays an average accuracy of 90.89% and a detection speed of 59.42 frame/s on the self-made vehicle dataset.Compared with the original SSD algorithm,the improved SSD algorithm increases the accuracy by 2.65 percentage points and speed by 17.41 frame/s.This algorithm can identify and locate the vehicles in images more quickly and accurately.  
摘要:  
SSD算法利用多尺度特征图进行分类和位置回归,检测小目标效果优于YOLO算法,但SSD算法在进行车辆检测时存在漏检问题。为此,提出一种改进SSD算法。为提取更多的车辆特征信息,设计改进Inception模块替代SSD网络中的Conv8、Conv9和Conv10层。将浅层特征的位置信息和深层特征的语义信息进行均衡化融合,构建多尺度特征融合均衡化网络,提高小目标车辆识别率。在特征提取层均引入SENet,对不同特征通道的重要性进行重标定以提高模型性能。实验结果表明,改进后SSD算法在自制的车辆数据集上平均精度为90.89%,检测速度达到59.42 frame/s,相比改进前的SSD算法,在精度和速度上分别提高2.65个百分点和17.41 frame/s,能够更快速、准确地对图像中的车辆进行识别和定位。  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
9  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
12  
ISSN:  
1000-3428  
入藏号:  
CSCD:7130685  
Indexed Date:  
2022-04-02  
Institution Address:  
广西大学电气工程学院, 南宁, 广西 530004, 中国  
Institution Address:  
School of Electrical Engineering, Guangxi University, Nanning, Guangxi 530004, China  
Institution:  
广西大学电气工程学院  
Institution:  
School of Electrical Engineering, Guangxi University

### 1059. SPFDNet: Water Extraction Method Based on Spatial Partition and Feature

摘要: Extracting water information from remote-sensing images is of great research significance for applications such as water resource protection and flood monitoring. Current water extraction methods aggregated richer multi-level features to enhance the output results. In fact, there is a difference in the requirements for the water body and the water boundary. Indiscriminate multi-feature fusion can lead to perturbation and competition of information between these two types of features during the optimization. Consequently, models cannot accurately locate the internal vacancies within the water body with the external boundary. Therefore, this paper proposes a water feature extraction network with spatial partitioning and feature decoupling. To ensure that the water features are extracted with deep semantic features and stable spatial information before decoupling, we first design a chunked multi-scale feature aggregation module (CMFAM) to construct a context path for obtaining deep semantic information. Then, an information interaction module (IIM) is designed to exchange information between two spatial paths with two fixed resolution intervals and the two paths through. During decoding, a feature decoupling module (FDM) is developed to utilize internal flow prediction to acquire the main body features, and erasing techniques are employed to obtain boundary features. Therefore, the deep features of the water body and the detailed boundary information are supplemented, strengthening the decoupled body and boundary features. Furthermore, the integrated expansion recoupling module (IERM) module is designed for the recoupling stage. The IERM expands the water body and boundary features using expansion and adaptively compensates the transition region between the water body and boundary through information guidance. Finally, multi-level constraints are combined to realize the supervision of the decoupled features. Thus, the water body and boundaries can be extracted more accurately. A comparative validation analysis is conducted on the public datasets, including the gaofen image dataset (GID) and the gaofen2020 challenge dataset (GF2020). By comparing with seven SOTAs, the results show that the proposed method achieves the best results, with IOUs of 91.22 and 78.93, especially in the localization of water bodies and boundaries. By applying the proposed method in different scenarios, the results show the stable capability of the proposed method for extracting water with various shapes and areas.  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2072-4292  
入藏号:  
WOS:001352101100001  
Indexed Date:  
2024-11-17  
Institution Address:  
Changjiang Water Resources Comm, Changjiang River Sci Res Inst, Wuhan 430010, Peoples R China  
Institution Address:  
Minist Water Resources, Key Lab River Regulat & Flood Control Middle & Low, Wuhan 430010, Peoples R China  
Institution Address:  
Yangtze Univ, Sch Geosci, Wuhan 430100, Peoples R China  
Institution Address:  
Hunan Inst Water Resources & Hydropower Res, Changsha 410007, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Hunan Inst Water Resources & Hydropower Res

### 1060. MSNet: Multi-Scale Network for Object Detection in Remote Sensing Images

摘要: Remote sensing object detection (RSOD) encounters challenges in effectively extracting features of small objects in remote sensing images (RSIs). To alleviate these problems, we proposed a Multi-Scale Network for Object Detection in Remote Sensing Images (MSNet) with multi-dimension feature information. Firstly, we design a Partial and Pointwise Convolution Extraction Module (P2CEM) 2 CEM) to capture feature of object in spatial and channel dimension simultaneously. Secondly, we design a Local and Global Information Fusion Module (LGIFM), designed local information stack and context modeling module to capture texture information and semantic information within the multi-scale feature maps respectively. Moreover, the LGIFM enhances the ability of representing features for small objects and objects within complex backgrounds by allocating weights between local and global information. Finally, we introduce Local and Global Information Fusion Pyramid (LGIFP). With the aid of the LGIFM, the LGIFP enhances the feature representation of small object information, which contributes to dense connection across the multi-scale feature maps. Extensive experiments validate that our proposed method outperforms state-of-the-art performance. Specifically, MSNet achieves mean average precision (mAP) scores of 75.3%, 93.39%, 96.00%, and 95.62% on the DIOR, HRRSD, NWPU VHR-10, and RSOD datasets, respectively.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:001315361300001  
Indexed Date:  
2024-09-23  
Institution Address:  
Changan Univ, Sch Informat Engn, Xian 710064, Peoples R China  
Institution Address:  
Changan Univ, Sch Data Sci & Artificial Intelligence, Xian 710064, Peoples R China  
Institution Address:  
Australian Natl Univ, Sch Comp, Canberra, ACT 2600, Australia  
Institution:  
Institution:  
Institution:

### 1061. ALF-YOLO: Enhanced YOLOv8 based on multiscale attention feature fusion

摘要: Ship detection plays a crucial role in ensuring maritime transportation and navigation safety. However, accurately detecting multiscale ships remains a challenge due to the diversity of ship categories and locations, as well as interference from complex environments. Object detectors based on the You Only Look Once (YOLO) framework have demonstrated remarkable accuracy in automatic ship detection. In this paper, we integrate the Asymptotic Feature Pyramid Network (AFPN), Large Selective Kernel Attention Mechanism (LSK), and the fourth detection head into YOLOv8, developing a novel ALF-YOLO architecture. ALF-YOLO utilizes AFPN to enrich feature representation by integrating multiscale high-level semantic features and spatial details. It also incorporates a large selective kernel attention mechanism that dynamically adjusts its large spatial receptive field to focus more on crucial ship features, eliminating interference from complex environmental factors to enhance discriminative feature representations of ships. Additionally, we investigate the impact of different attention mechanisms on ship detection accuracy. Experimental results indicate that by integrating the outputs of several modules, our proposed ALF-YOLO model improves the classification and localization capability of targets at each stage. Compared to YOLOv8, ALF-YOLO achieved a relative increase of 0.41% and 0.43% in mAP@0.50 on the Seaships and McShips datasets, respectively. Across different evaluation criteria, the overall performance of the ALF-YOLO method surpasses existing ship detection methods.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
12  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
13  
ISSN:  
0029-8018  
eISSN:  
1873-5258  
入藏号:  
WOS:001247175500001  
Indexed Date:  
2024-06-21  
Institution Address:  
Dalian Maritime Univ, Nav Coll, Dalian 116026, Liaoning, Peoples R China  
Institution Address:  
Dalian Maritime Univ, Informat Sci & Technol Coll, Dalian 116026, Liaoning, Peoples R China  
Institution:  
Institution:

### 1062. Improved heterogeneous data fusion and multi-scale feature selection

摘要: The diagnosis of the disease requires a variety of data and indicators. In order to make the computer perform intelligent computation and diagnosis from a doctor's perspective, complementary information between different modal data needs to be taken into account. Meanwhile, the redundancy features with key information should be selected in order to reduce the complexity of calculation. In this study, an adaptive dynamic loss function is proposed to weight different scales in the multi-scale expansion network of pathological images according to the doctor's diagnosis process. And an ant colony algorithm based on maximum information coefficient correlation was designed for unsupervised feature selection of fusion features combined image feature and patient differential genes. Experimental results show that the addition of pathological image information and genetic information plays an important role in the classification of lung cancer subtypes. Compared with other feature selection methods, the proposed algorithm can quickly converge. Combining pathological image and gene expression matrix for cancer diagnosis can improve the diagnostic accuracy of specific patients, with an accuracy of 95.62 and AUC achieves 0.897. The proposed method has high effectiveness and superior performance in the classification of lung cancer.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1532-0626  
eISSN:  
1532-0634  
入藏号:  
WOS:000680721400001  
Indexed Date:  
2021-08-08  
Institution Address:  
Taiyuan Univ Technol, Coll Informat & Comp, Taiyuan, Peoples R China  
Institution Address:  
Shanxi Canc Hosp, Dept Radiol Imaging, Taiyuan, Peoples R China  
Institution Address:  
Shanxi Prov Peoples Hosp, Clin Lab, Taiyuan, Peoples R China  
Institution:  
Institution:  
Shanxi Canc Hosp  
Institution:

### 1063. Research on Detecting Bearing-Cover Defects Based on Improved YOLOv3

摘要: Detecting defects, which is a branch of target detection in the field of computer vision, is widely used in factory production. To solve the problems in existing detection algorithms that relate to their insensitivity to large or medium defect targets on bearing covers, their difficulty in detecting subtle defects effectively and their lack of real-time detection, in this work, we establish a large-scale bearing-cover defect dataset and propose an improved YOLOv3 network model. The proposed model is divided into four submodels: the bottleneck attention network (BNA-Net), the attention prediction subnet model, the defect localization subnet model, and the large-size output feature branch. To test the generality, robustness and practicability of the new model, we design a comparative experiment under abnormal illumination conditions. We design an ablation experiment to verify the validity of the proposed submodules. The experimental results show that our model solves the problem of the YOLOv3 algorithm's insensitivity to medium or large targets and satisfies real-time detection conditions. The mAP result is 69.74%, which is 16.31%, 13.4%, 13%, 10.9%, and 7.2% more than that of YOLOv3, EfficientDet-D2, YOLOv5, YOLOv4, and PP-YOLO, respectively.  
Web of Science 核心合集中的 "被引频次":  
34  
在中国科学引文数据库中的被引频次:  
9  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
43  
ISSN:  
2169-3536  
入藏号:  
WOS:000609793000001  
Indexed Date:  
2021-02-08  
Institution Address:  
Univ Sci & Technol Liaoning, Sch Comp Sci & Software Engn, Anshan 114051, Peoples R China  
Institution:

### 1064. Method for enhancement of the multi-scale low-light image by combining

摘要: The low light environment affects the image capture equipment,resulting in low contrast,low brightness,and difficulty in distinguishing objects.In order to improve image quality,a method for enhancement of the multi-scale low-light image by combining an attention guidance is proposed.First,a dense residual network is constructed as a multi-scale feature extractor to extract feature maps at different scales in low light images,and the extracted feature maps are fused by using a modified RefineNet,which makes full use of the feature information in the image.Meanwhile,an interpretable attention mechanism is designed to generate an attention graph based on the results of edge detection.Then by combining a loss function the network is guided through training.The purpose is to enhance edge detail information hidden in the dark without increasing the network's inference burden.Finally,experiments are completed on synthetic images and SID (See-in-the-Dark)datasets,with the results showing that the proposed method can effectively improve brightness and contrast,restore image edge details as well as improve subjective visual effects.Compared to the contrast algorithm,the PSNR and SSIM are improved by at least 0.79dB and 0.119 on average,respectively.  
摘要:  
弱光环境导致图像采集设备拍摄的照片呈现出对比度低、亮度较暗、目标物难以分辨等特点。为了改善图像质量,提出了一种融合注意力引导的多尺度低照度图像增强方法。首先,构建密集残差网络作为多尺度特征提取器,用于提取低照度图像中不同尺度的特征图;其次,利用改进的RefineNet对提取出的不同尺度的特征图进行融合,以便充分利用图像中的特征信息;同时,在网络中引入注意力机制,基于边缘检测结果生成注意力图,并与损失函数相结合来引导网络进行训练,在不增加网络推理负担的同时,增强隐藏在黑暗中的细节信息;最后,实验分别选用合成图像和SID(See-in-the-Dark)数据集进行训练与测试。相较于对比算法,峰值信噪比(PSNR)和结构相似性(SSIM)分别平均提高了约0.79dB和0.119。结果表明,所提方法能有效提高亮度和对比度,恢复图像边缘细节,主观视觉效果得到提升。  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
2  
ISSN:  
1001-2400  
入藏号:  
CSCD:7439967  
Indexed Date:  
2023-07-14  
Institution Address:  
天津大学微电子学院, 天津 300072, 中国  
Institution Address:  
School of Microelectronics,Tianjin University, Tianjin 300072, China  
Institution:  
天津大学微电子学院  
Institution:  
School of Microelectronics,Tianjin University

### 1065. GLER-BiGRUnet: A Surface Deformation Prediction Model Fusing Multiscale

摘要: Accurate surface deformation (SD) predictions are critical for early warning and timely remediation of infrastructure damage. However, the current SD prediction models do not integrate the multiscale features of InSAR SD and environmental factors (EFs), which make their prediction results inaccurate. To address these limitations, we proposed a bidirectional gated recurrent unit (BiGRU) multioutput SD prediction network (GLER-BiGRUnet), which mainly included global-local feature extraction (GLFE), multifactor cross-attention residual (MCAR), and local residual module embedded in self-attention mechanism (RCSA) modules. Specifically, dense and one-dimensional convolutional layers were concatenated in the GLFE module to extract global-local SD features. The long time-series dependence between EFs and SD was learned in the MCAR module using the multihead cross-attention mechanism to obtain the corresponding attention weight feature matrix. The residual connection and self-attention mechanisms were used in the RCSA module to merge the multiscale features and enhance the model fitting ability. We chose four typical regions in the permafrost area of Qinghai-Tibet Railway as the scene for the experiment. The spatial distribution and local profile exhibited relatively small discrepancies between the prediction results of the GLER-BiGRUnet model and the InSAR SD. Meanwhile, the average root-mean-square error of the GLER-BiGRUnet model in the four typical regions was 0.19 mm, and the proposed model had the best evaluation index compared with other SD prediction models. Additionally, the prediction trend of SD of the proposed GLER-BiGRUnet model was consistent with the original InSAR SD, and the prediction results were more stable than those of the other prediction models. The SD prediction model proposed in this article contributes to early warning of SD.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001308275000025  
Indexed Date:  
2024-09-18  
Institution Address:  
Lanzhou Jiaotong Univ, Fac Geomat, Lanzhou 730070, Peoples R China  
Institution Address:  
Natl Local Joint Engn Res Ctr Technol & Applicat N, Lanzhou 730070, Peoples R China  
Institution Address:  
Gansu Prov Key Lab Sci & Technol Surveying & Mappi, Lanzhou 730070, Peoples R China  
Institution:  
Institution:  
Natl Local Joint Engn Res Ctr Technol & Applicat N  
Institution:  
Gansu Prov Key Lab Sci & Technol Surveying & Mappi

### 1066. Learn Robust Pedestrian Representation Within Minimal Modality

摘要: Visible-infrared person re-identification has attracted extensive attention from the community due to its potential great application prospects in video surveillance. There are huge modality discrepancies between visible and infrared images caused by different imaging mechanisms. Existing studies alleviate modality discrepancies by aligning modality distribution or extracting modality-shared features on the original image. However, they ignore a key solution, i.e., converting visible images to gray images directly, which is efficient and effective to reduce modality discrepancies. In this paper, we transform the cross-modality person re-identification task from visible-infrared images to gray-infrared images, which is named as the minimal modality discrepancy. In addition, we propose a pyramid feature integration network (PFINet) which mines the discriminative refined features of pedestrian images and fuses high-level and semantically strong features to build a robust pedestrian representation. Specifically, PFINet first performs the feature extraction from concrete to abstract and the top-down semantic transfer to obtain multi-scale feature maps. Second, the multi-scale feature maps are inputted to the discriminative-region response module to emphasize the identity-discriminative regions by the spatial attention mechanism. Finally, the pedestrian representation is obtained by the feature integration. Extensive experiments demonstrate the effectiveness of PFINet which achieves the rank-1 accuracy of 81.95% and mAP of 74.49% on the multi-all evaluation mode of the SYSU-MM01 dataset.  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
4  
ISSN:  
1000-9000  
eISSN:  
1860-4749  
入藏号:  
WOS:000812520400011  
Indexed Date:  
2022-06-29  
Institution Address:  
China Univ Petr East China, Coll Comp Sci & Technol, Qingdao 266580, Peoples R China  
Institution:

### 1067. Cascaded attention-induced difference representation learning for

摘要: Change Detection (CD) is the process of recognizing and quantitatively evaluating changes in surface objects at the exact location but at different times from remote sensing images that may be caused by extreme heat events. Attention-based CD methods have gained much traction because of their ability to concentrate on change regions. However, most current attention-based methods only calculate the attention matrix based on features extracted from a single image but fail to consider the correlation of features extracted from images at different times. The effectiveness of the correlation of learned features from change regions in bi-temporal images is an essential element influencing the improvement of CD performance. This paper proposes a Cascaded Attention-Induced Difference Representation Learning (CADRL) method for multispectral CD to explore the correlation of features extracted from bi-temporal images to obtain more discriminative features. The proposed CADRL method contains three modules: the feature extraction module, the Cascaded Cross-attention based Difference Learning Module (CCADLM) and the detection module. First, the feature extraction module extracts multi-scale features from bi-temporal images. Then, CCADLM generates more discriminative features by fusing difference features and the cross-attention matrix learned from the temporal attention unit from multiple levels to explore the correlation information of change regions in bi-temporal images. Finally, the learned discriminative features are fed to the detection module to gain the final detection map. Experimental results on three multispectral datasets demonstrate that the CADRL method outperforms other existing CD algorithms.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
1569-8432  
eISSN:  
1872-826X  
入藏号:  
WOS:001011200600001  
Indexed Date:  
2023-07-01  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Comp Sci & Technol, Shaanxi Key Lab Network Data Anal & Intelligent Pr, Xian 710121, Shaanxi, Peoples R China  
Institution Address:  
Fuzhou Univ, Coll Phys & Informat Engn, Fuzhou, Peoples R China  
Institution:  
Institution:

### 1068. Three-dimensional visualization of thyroid ultrasound images based on

摘要: BackgroundUltrasound three-dimensional visualization, a cutting-edge technology in medical imaging, enhances diagnostic accuracy by providing a more comprehensive and readable portrayal of anatomical structures compared to traditional two-dimensional ultrasound. Crucial to this visualization is the segmentation of multiple targets. However, challenges like noise interference, inaccurate boundaries, and difficulties in segmenting small structures exist in the multi-target segmentation of ultrasound images. This study, using neck ultrasound images, concentrates on researching multi-target segmentation methods for the thyroid and surrounding tissues.MethodWe improved the Unet++ to propose PA-Unet++ to enhance the multi-target segmentation accuracy of the thyroid and its surrounding tissues by addressing ultrasound noise interference. This involves integrating multi-scale feature information using a pyramid pooling module to facilitate segmentation of structures of various sizes. Additionally, an attention gate mechanism is applied to each decoding layer to progressively highlight target tissues and suppress the impact of background pixels.ResultsVideo data obtained from 2D ultrasound thyroid serial scans served as the dataset for this paper.4600 images containing 23,000 annotated regions were divided into training and test sets at a ratio of 9:1, the results showed that: compared with the results of U-net++, the Dice of our model increased from 78.78% to 81.88% (+ 3.10%), the mIOU increased from 73.44% to 80.35% (+ 6.91%), and the PA index increased from 92.95% to 94.79% (+ 1.84%).ConclusionsAccurate segmentation is fundamental for various clinical applications, including disease diagnosis, treatment planning, and monitoring. This study will have a positive impact on the improvement of 3D visualization capabilities and clinical decision-making and research in the context of ultrasound image.  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
1475-925X  
入藏号:  
WOS:001182105000001  
PubMed ID:  
38468262  
Indexed Date:  
2024-03-18  
Institution Address:  
Sichuan Univ, Coll Biomed Engn, Chengdu, Sichuan, Peoples R China  
Institution Address:  
China Japan Friendship Hosp, Beijing, Peoples R China  
Institution Address:  
Gen Hosp Western Theater Command, Dept Ultrasound, Chengdu, Sichuan, Peoples R China  
Institution Address:  
Sichuan Univ, Dept Ultrasound, West China Hosp, Chengdu, Sichuan, Peoples R China  
Institution Address:  
Highong Intellimage Med Technol Tianjin Co Ltd, Tianjin, Peoples R China  
Institution:  
Institution:  
Institution:  
Gen Hosp Western Theater Command  
Institution:  
Institution:  
Highong Intellimage Med Technol Tianjin Co Ltd

### 1069. A multiscale dilated attention network for hyperspectral image

摘要: Hyperspectral imaging is an image obtained by combining spectral detection technology and imaging technology, which can collect electromagnetic spectra in the wavelength range of visible light to near-infrared. It is an important research content in the field of ground observation in hyperspectral remote sensing. However, hyperspectral image face significant challenges in classification task due to their high spectral dimensions, lack of labeled samples, and strong correlation between bands. In order to fully extract features from both spectral and spatial dimensions and improve classification accuracy in the case of limited training samples, a multiscale dilated attention network is proposed for hyperspectral image classification. First, a three-dimensional convolutional layer is used to extract the shallow features of the image. Then, a multiscale dilated attention module is proposed by combining dilated convolution and channel attention. Using ordinary convolution and dilated convolution to form different receptive fields. Channel attention is used to remodel the obtained multiscale features, enhancing the inter-channel correlation. After that, a multiscale spatial-spectral attention module is constructed using multiple asymmetric convolutions to obtain spatial and spectral attention features at different positions, further enhancing important feature suppression over non-important features. Finally, using softmax to classify the obtained features. Using Indian Pines, Pavia University, KSC and University of Houston as experimental datasets, the overall classification accuracy of this paper's method achieved 98.97%, 99.14%, 99.45%, and 98.56% respectively, using only 5%, 1%, 10%, and 10% of training samples per class. Compared with seven advanced classification methods, the experimental results show that the proposed method can achieve the highest classification accuracy with limited training samples. (c) 2024 COSPAR. Published by Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0273-1177  
eISSN:  
1879-1948  
入藏号:  
WOS:001359172000001  
Indexed Date:  
2024-11-26  
Institution Address:  
Liaoning Tech Univ, Sch Geomat, Fuxin 123000, Peoples R China  
Institution Address:  
Liaoning Tech Univ, Sch Software, Huludao 125105, Peoples R China  
Institution:  
Institution:

### 1070. An improved YOLOv5 method for large objects detection with multi-scale

摘要: SSD and YOLOv5 are the one-stage object detector representative algorithms. An improved one-stage object de-tector based on the YOLOv5 method is proposed in this paper, named Multi-scale Feature Cross-layer Fusion Net-work (M-FCFN). Firstly, we extract shallow features and deep features from the PANet structure for cross-layer fusion and obtain a feature scale different from 80 x 80, 40 x 40, and 20 x 20 as output. Then, according to the single shot multi-box detector, we propose the different scale features which are obtained by cross-layer fusion for dimension reduction and use it as another output for prediction. Therefore, two completely different feature scales are added as the output. Features of different scales are necessary for detecting objects of different sizes, which can increase the probability of object detection and significantly improve detection accuracy. Finally, aiming at the Autoanchor mechanism proposed by YOLOv5, we propose an EIOU k-means calculation. We have compared the four model structures of S, M, L, and X of YOLOv5 respectively. The problem of missed and false de-tections for large objects is improved which has better detection results. The experimental results show that our methods achieve 89.1% and 67.8% mAP@0.5 on the PASCAL VOC and MS COCO datasets. Compared with the YOLOv5\_S, our methods improve by 4.4% and 1.4% mAP@ [0.5:0.95] on the PASCAL VOC and MS COCO datasets. Compared with the four models of YOLOv5, our methods have better detection accuracy for large objects. It should be more attention that our method on the large-scale mAP@ [0.5:0.95] is 5.4% higher than YOLOv5\_S on the MS COCO datasets. (c) 2022 Elsevier B.V. All rights reserved.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
40  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
7  
被引频次合计:  
41  
ISSN:  
0262-8856  
eISSN:  
1872-8138  
入藏号:  
WOS:000843898700004  
Indexed Date:  
2022-08-30  
Institution Address:  
Chongqing Univ Posts & Telecommun, Sch Software Engn, Chongqing 400065, Peoples R China  
Institution Address:  
Southwest Comp Co Ltd, Inst Informat Technol, Chongqing 400060, Peoples R China  
Institution:  
Institution:  
Southwest Comp Co Ltd

### 1071. A lightweight fine-grained pelagic fish recognition algorithm based on

摘要: In order to enhance the fine-grained recognition of fish species, this paper proposes a lightweight object detection model YOLOv8n-DFG. The model accurately identifies six pelagic fish species including flatfin sailfish, striped marlin, Atlantic bluefin tuna, bigeye tuna, longtail tuna, and yellowfin tuna, meeting the requirements for real-time detection and fishing monitoring. Firstly, by introducing FasterNet Block and EMA attention mechanism into the YOLOv8 network structure to improve C2f and obtain the C2f-FE module, this model enhances feature extraction accuracy and operational efficiency. Subsequently, it combines BiFPN structure with C2f-FE module to construct a fast and lightweight neck network structure that achieves multi-scale feature fusion. Additionally, Dysample dynamic upsampling module is introduced along with porting of ADown downsampling module from YOLOv9 to optimize feature pyramid sampling method named as YOLOv8-FG. Finally, using large-sized YOLOv8s-FG as teacher network and small-sized YOLOv8n-FG as student network based on CWD loss intermediate layer feature distillation method constructs the final model YOLOv8n-DFG. Experimental results on a dataset containing six morphologically similar fish species demonstrate the effectiveness of these improvements and distillation effects are significant. Compared to YOLOv8n, precision has increased by 7.8%, recall by 3.3%, mAP@0.5 by 5.6%, mAP@0.5:0.95 by 6.7%, while FlOPs decreased by 42% with a reduction in model size of 58%. The results indicate that our proposed YOLOv8n-DFG demonstrates exceptional accuracy and real-time performance, effectively fulfilling the requirements for real-time fine-grained fish recognition.  
Um die feink & ouml;rnige Erkennung von Fischarten zu verbessern, schl & auml;gt dieser Artikel ein leichtgewichtiges Zielerkennungsmodell namens YOLOv8n-DFG vor. Das Modell erkennt pr & auml;zise sechs pelagische Fischarten, darunter Istiophorus platypterus, Tetrapturus auda, Thunnus thynnus, Thunnus obesus, Thunnus alalunga und Thunnus albacares, und erf & uuml;llt damit die Anforderungen f & uuml;r die Echtzeit-Erkennung und Fischerei & uuml;berwachung. Zun & auml;chst wird der FasterNet-Block sowie der EMA-Aufmerksamkeitsmechanismus in die YOLOv8-Netzwerkstruktur eingef & uuml;hrt, um das C2f zu verbessern und das C2f-FE-Modul zu erhalten. Dadurch wird die Genauigkeit der Merkmalsextraktion und die Betriebseffizienz erh & ouml;ht. Anschlie ss end wird die BiFPN-Struktur mit dem C2f-FE-Modul kombiniert, um eine schnelle und leichtgewichtige Nacken-Netzwerkstruktur zu konstruieren, die eine mehrskalige Merkmalsfusion erm & ouml;glicht. Zus & auml;tzlich wird das Dysample-Dynamische Upsampling-Modul eingef & uuml;hrt, zusammen mit der Portierung des ADown-Downsampling-Moduls aus YOLOv9, um die Merkmals-Pyramiden-Sampling-Methode zu optimieren, die als YOLOv8-FG bezeichnet wird. Schlie ss lich wird unter Verwendung des gro ss formatigen YOLOv8s-FG als Lehrer-Netzwerk und des kleinformatigen YOLOv8n-FG als Sch & uuml;ler-Netzwerk basierend auf der CWD-Verlustmethoden zur Zwischenschicht-Feature-Distillation das endg & uuml;ltige Modell YOLOv8n-DFG konstruiert. Die experimentellen Ergebnisse auf einem Datensatz mit sechs morphologisch & auml;hnlichen Fischarten zeigen die Wirksamkeit dieser Verbesserungen, und die Distillations-Effekte sind signifikant. Im Vergleich zu YOLOv8n wurden die Pr & auml;zision um 7,8 %, der Recall um 3,3 %, der mAP@0.5 um 5,6 %, der mAP@0.5:0.95 um 6,7 % erh & ouml;ht, w & auml;hrend die FlOPs um 42 % reduziert und die Modellgr & ouml;ss e um 58 % verringert wurden. Die Ergebnisse zeigen, dass das vorgeschlagene YOLOv8n-DFG au ss ergew & ouml;hnliche Genauigkeit und Echtzeit-Leistung bietet und die Anforderungen f & uuml;r die Echtzeit-Erkennung von feingliedrigen Fischen effektiv erf & uuml;llt.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0967-6120  
eISSN:  
1573-143X  
入藏号:  
WOS:001390043600020  
Indexed Date:  
2025-01-11  
Institution Address:  
Shanghai Ocean Univ, Coll Engn Sci & Technol, Shanghai 201306, Peoples R China  
Institution Address:  
East China Univ Sci & Technol, Sch Mech & Power Engn, Shanghai 200237, Peoples R China  
Institution Address:  
Shanghai Ocean Univ, Coll Marine Sci, Shanghai 201306, Peoples R China  
Institution Address:  
Shanghai Engn Res Ctr Marine Renewable Energy, Shanghai 201306, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Shanghai Engn Res Ctr Marine Renewable Energy

### 1072. Deep co-supervision and attention fusion strategy for automatic COVID-19

摘要: Due to the irregular shapes,various sizes and indistinguishable boundaries between the normal and infected tissues, it is still a challenging task to accurately segment the infected lesions of COVID-19 on CT images. In this paper, a novel segmentation scheme is proposed for the infections of COVID-19 by enhancing supervised information and fusing multi-scale feature maps of different levels based on the encoder-decoder architecture. To this end, a deep collaborative supervision (Co-supervision) scheme is proposed to guide the network learning the features of edges and semantics. More specifically, an Edge Supervised Module (ESM) is firstly designed to highlight low-level boundary features by incorporating the edge supervised information into the initial stage of down-sampling. Meanwhile, an Auxiliary Semantic Supervised Module (ASSM) is proposed to strengthen high-level semantic information by integrating mask supervised information into the later stage. Then an Attention Fusion Module (AFM) is developed to fuse multiple scale feature maps of different levels by using an attention mechanism to reduce the semantic gaps between high-level and low-level feature maps. Finally, the effectiveness of the proposed scheme is demonstrated on four various COVID-19 CT datasets. The results show that the proposed three modules are all promising. Based on the baseline (ResUnet), using ESM, ASSM, or AFM alone can respectively increase Dice metric by 1.12%, 1.95%,1.63% in our dataset, while the integration by incorporating three models together can rise 3.97%. Compared with the existing approaches in various datasets, the proposed method can obtain better segmentation performance in some main metrics, and can achieve the best generalization and comprehensive performance. (c) 2021 Elsevier Ltd. All rights reserved.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
26  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
26  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:000736980400009  
PubMed ID:  
34848897  
Indexed Date:  
2021-12-03  
Institution Address:  
Zhejiang Univ Technol, Coll Comp Sci & Technol, Hangzhou 310023, Peoples R China  
Institution Address:  
Key Lab Visual Media Intelligent Proc Technol Zhe, Hangzhou 310023, Peoples R China  
Institution Address:  
Univ Rouen Normandy, LITIS EA 4108, F-76183 Rouen, France  
Institution:  
Institution:  
Key Lab Visual Media Intelligent Proc Technol Zhe  
Institution:

### 1073. Spectral Energy Model-Driven Inversion of XCO<sub>2</sub> in IPDA Lidar

摘要: Carbon observation satellites based on passive theory (e.g., OCO-2/3, GOSAT-1/2, and TanSat) have relatively high carbon dioxide column concentration (XCO2) accuracy when the observation conditions are met. Passive satellites have data bias and coverage deficiencies due to cloud cover, low albedo, low-light conditions, and aerosol scattering, resulting in carbon observation satellites based on passive theory that cannot meet the demand for high-precision, all-day, all-weather XCO2 monitoring. Active detection satellites are urgently needed to support global carbon sources, sinks, and carbon neutrality. China intends to launch a sensor satellite with active detection of XCO2 in the coming years. In this work, based on the satellite's scaled-down airborne experiments, a spectral energy model was developed to optimize the conventional inversion algorithm and achieve a more accurate XCO2 inversion. The 1.572-mu m integrated path differential absorption (IPDA) lidar column length is used indirectly to evaluate the accuracy of the spectral energy model for signal extraction. Also, the experimental results show that the accuracy of the signal extracted by the 1.572-mu m IPDA lidar column length is 0.74 and 6.20 m at sea and on land based on the indirect evaluation of the length of the 1.572-mu m IPDA lidar column length. The optimized XCO2 was evaluated (standard deviation as an evaluation metric) and its XCO2 standard deviation reduced by 31%, 63%, and 66% in the ocean, plains, and mountains, respectively. Our algorithm can obtain the XCO2 with a consistent trend by using XCO2 from the OCO-2 satellite as a reference. The calculated XCO2 is more accurate in areas dominated by anthropogenic factors (plains), due to the accuracy of the IPDA detection mechanism. This algorithm improves the accuracy and robustness of XCO2 inversion and has important reference significance for the IPDA lidar carried by China's satellites to be launched in this year.  
在 BIOSIS Citation Index 中的被引频次:  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
44  
被引频次合计:  
46  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000960565400006  
Indexed Date:  
2023-04-30  
Institution Address:  
Wuhan Univ, Sch Elect Informat, Wuhan 473072, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Remote Sensing & Informat Engn, Wuhan 430079, Peoples R China  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & R, Wuhan 430079, Peoples R China  
Institution:  
Institution:  
Institution:

### 1074. Real-time Infrared Target Detection Algorithm for Embedded System in

摘要: In order to solve the problems of low accuracy and recall rate of infrared target detection under complex background conditions, as well as slow inference speed of network model on embedded computing platform, lightweight network YOLOv4-Tiny was taken as the basic architecture of the algorithm, combined with visual attention mechanism and spatial pyramid pooling idea. Two real-time infrared target detection networks for embedded systems are proposed. Among them, there are a lot of background interference information in target detection in infrared complex scenes. Therefore, the visual attention mechanism is used to effectively learn the weight distribution of the feature map, recalibrate the feature map, strengthen the focus on the target, reduce the influence of irrelevant background information and improve the detection and recognition ability of the model. Spatial pyramid pooling can fuse multi-scale features, enrich the information of feature maps and improve the ability of infrared target recognition and location at different scales. Grad-CAM was used to visualize the feature map strengthened by the attention mechanism, showing the attention of the network model to the target region. The training is carried out on a 2080Ti GPU computer platform using the transfer learning strategy, and deployed on the Atlas 200 DK embedded computing platform with Ascend 310 AI chip as the core. The experimental results show that compared with the original network YOLOv4-Tiny, the infrared images with a resolution of 640 pixels x 512 pixels are detected on the computer platform. The average accuracy and recall rate of the proposed YOLOv4-Tiny+SE+SPP network were improved by 13.96% and 20.14%, respectively, and the inference speed reached 212 FPS. The average accuracy and recall rate of the proposed YOLOv4-Tiny+CBAM+SPP network were improved by 15.75% and 22.41%, respectively, and the inference speed reached 202 FPS. On Atlas 200 DK embedded computing platform, infrared images with a resolution of 640 pixelx512 pixel are detected, compared with the original network YOLOv4-Tiny. The average accuracy and recall rate of the proposed YOLOv4-Tiny+SE+SPP network were improved by 12.36% and 18.6%, respectively, and the inference speed reached 78 FPS. The average accuracy and recall rate of the proposed network YOLOv4-Tiny+CBAM+SPP are improved by 15.94% and 22.89%, respectively, and the inference speed reaches 71 FPS, which can meet the needs of real-time detection and tracking of infrared targets in military and security fields.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1004-4213  
入藏号:  
WOS:000808045500002  
Indexed Date:  
2022-07-01  
Institution Address:  
Nanjing Univ Sci & Technol, Sch Mech Engn, Nanjing 210094, Peoples R China  
Institution Address:  
Nanjing Plandech Intelligent Technol Co LTD, Nanjing 210014, Peoples R China  
Institution Address:  
Shanghai Univ, Inst Artificial Intelligence, Shanghai 200444, Peoples R China  
Institution:  
Institution:  
Nanjing Plandech Intelligent Technol Co LTD  
Institution:

## Convolutional Neural Network (CNN)

### 1. Traffic Sign Detection Based on Lightweight Multiscale Feature Fusion

摘要: Traffic sign detection is a research hotspot in advanced assisted driving systems, given the complex background, light transformation, and scale changes of traffic sign targets, as well as the problems of slow result acquisition and low accuracy of existing detection methods. To solve the above problems, this paper proposes a traffic sign detection method based on a lightweight multiscale feature fusion network. Since a lightweight network model is simple and has fewer parameters, it can greatly improve the detection speed of a target. To learn more target features and improve the generalization ability of the model, a multiscale feature fusion method can be used to improve recognition accuracy during training. Firstly, MobileNetV3 was selected as the backbone network, a new spatial attention mechanism was introduced, and a spatial attention branch and a channel attention branch were constructed to obtain a mixed attention weight map. Secondly, a feature-interleaving module was constructed to convert the single-scale feature map of the specified layer into a multiscale feature fusion map to realize the combined encoding of high-level semantic information and low-level semantic information. Then, a feature extraction base network for lightweight multiscale feature fusion with an attention mechanism based on the above steps was constructed. Finally, a key-point detection network was constructed to output the location information, bias information, and category probability of the center points of traffic signs to achieve the detection and recognition of traffic signs. The model was trained, validated, and tested using TT100K datasets, and the detection accuracy of 36 common categories of traffic signs reached more than 85%, among which the detection accuracy of five categories exceeded 95%. The results showed that, compared with the traditional methods of Faster R-CNN, CornerNet, and CenterNet, traffic sign detection based on a lightweight multiscale feature fusion network had obvious advantages in the speed and accuracy of recognition, significantly improved the detection performance for small targets, and achieved a better real-time performance.

### 2. Capsule Network With Multiscale Feature Fusion for Hidden Human Activity

摘要: This article considers the problem of human activity classification behind the walls using ultrawideband (UWB) radar. The complex-valued multiscale feature fusion capsule network (CV-MCNet) is proposed, which consists of a feature extractor, a multiscale feature fusion (MFF) block, and a capsule block. Specifically, the feature extractor with two complex-valued convolutional layers is designed to extract the deep features from the range profiles. Then, the MFF block is developed to enrich the feature representation of the activity. Finally, a capsule block is applied to implicitly encode the spatial relationship among the features in vector form and aggregate the vectors to get accurate classification results. The proposed CV-MCNet is evaluated by real data, and the results show that it achieves better classification performance compared with the deep convolutional neural network (DCNN), convolutional autoencoder (CAE), and complex-valued convolutional neural network (CV-CNN).

### 3. Lightweight Channel Attention and Multiscale Feature Fusion

摘要: High-resolution remote sensing image scene classification has attracted widespread attention as a basic earth observation task. Remote sensing scene classification aims to assign specific semantic labels to remote sensing scene images to serve specified applications. Convolutional neural networks are widely used for remote sensing image classification due to their powerful feature extraction capabilities. However, the existing methods have not overcome the difficulties of large-scene remote sensing images of large intraclass diversity and high interclass similarity, resulting in low performance. Therefore, we propose a new remote sensing scene classification method that combines lightweight channel attention and multiscale feature fusion discrimination, called LmNet. First, ResNeXt is used as the backbone; second, a new lightweight channel attention mechanism is constructed to quickly and adaptively learn the salient features of important channels. Furthermore, we designed a multiscale feature fusion discrimination framework, which fully integrates shallow edge feature information and deep semantic information to enhance feature representation capabilities and uses multiscale features for joint discrimination. Finally, a cross-entropy loss function based on label smoothing is built to reduce the influence of interclass similarity on feature representation. In particular, our lightweight channel attention and multiscale feature fusion mechanism can be flexibly embedded in any advanced backbone as a functional module. The experimental results on three large-scale remote sensing scene classification datasets show that compared with the existing advanced methods, our proposed high-efficiency end-to-end scene classification method has reached state-of-the-art. Moreover, our method has a weaker dependence on labeled data and provided better generalization performance.

### 4. Remote Sensing Image Semantic Segmentation Method Based on a Deep

摘要: There are many problems with remote sensing images, such as large data scales, complex illumination conditions, occlusion, and dense targets. The existing semantic segmentation methods for remote sensing images are not accurate enough for small and irregular target segmentation results, and the edge extraction results are poor. The authors propose a remote sensing image segmentation method based on a DCNN and multiscale feature fusion. Firstly, an end-to-end remote sensing image segmentation model using complete residual connection and multiscale feature fusion was designed based on a deep convolutional encoder-decoder network. Secondly, weighted high-level features were obtained using an attention mechanism, which better preserved the edges, texture, and other information of remote sensing images. The experimental results on ISPRS Potsdam and Urban Drone datasets show that compared with the comparison methods, this method has better segmentation effect on small and irregular objects and achieves the best segmentation performance while ensuring the computation speed.

### 5. RaHFF-Net: Recall-Adjustable Hierarchical Feature Fusion Network for

摘要: Remote sensing (RS) image change detection (CD) aims to identify areas of interest that have changed between bitemporal images. For complex scenarios (e.g., varying lighting conditions), the diverse shapes and scales of the changed areas is especially vulnerable to cause CD models to suffer from serious missed detections. To address aforementioned problem, we propose a high recall multiscale feature fusion model for RS change interpretation. Initially, the RaHFF-Net extracts hierarchical multiscale feature from bitemporal RS images; Then, it employs CNN and Transformer to effectively merge local and global information across same-scale, cross-scale, and multiscale features. Finally, to address the issue of instance imbalance in CD, a novel hyperexpectation push pull loss regularization term is proposed. This loss function is designed to elevate the expected predictions of positive instances across the dataset, thereby enabling the development of a deep learning model with a high recall rate.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001409622100016  
Indexed Date:  
2025-02-09  
Institution Address:  
North Univ China, Dept Comp Sci & Technol, Taiyuan 030051, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Geog Sci & Nat Resources Res, State Key Lab Resources & Environm Informat Syst, Beijing 100101, Peoples R China  
Institution:  
Institution:

### 6. Efficient Multi-Scale Feature Fusion for Image Manipulation Detection

摘要: Convolutional Neural Network (CNN) has made extraordinary progress in image classification tasks. However, it is less effective to use CNN directly to detect image manipulation. To address this problem, we propose an image filtering layer and a multi-scale feature fusion module which can guide the model more accurately and effectively to perform image manipulation detection. Through a series of experiments, it is shown that our model achieves improvements on image manipulation detection compared with the previous researches.  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0916-8532  
eISSN:  
1745-1361  
入藏号:  
WOS:000790987400028  
Indexed Date:  
2022-05-14  
Institution Address:  
Shanghai Univ, Sch Commun & Informat Engn, Shanghai, Peoples R China  
Institution:

### 7. Multiscale feature fusion for malicious HTTP request detection

摘要: This paper proposed a multiscale feature fusion approach for malicious HTTP request detection.Firstly,it modeled the HTTP request in both word-level and character-level.Secondly,it extracted the high level sematic information in HTTP request by using a specially designed convolutional neural network(CNN).Thirdly,it jointly learnt the multiscale representation for HTTP request with the help of multimodal learning techniques.Finally,it adopted a linear classifier for classification.Experiments on public HTTP CSIC 2010 dataset and WAF dataset show that the proposed method has large improvement on the performance against existing state-of-the-art methods.  
摘要:  
针对当前网络环境中恶意HTTP请求攻击泛滥的问题,提出了一种多尺度特征融合的检测方法。首先从单词级和字符级两个尺度对HTTP请求进行建模,然后使用卷积神经网络提取其高阶语义特征;再借助多尺度特征融合技术,学习HTTP请求的多尺度公共向量表示;最后使用线性分类器进行分类。实验结果表明该方法性能在HTTP CSIC 2010数据集和WAF真实数据集上优于现有方法。  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1001-3695  
入藏号:  
CSCD:6934441  
Indexed Date:  
2021-06-18  
Institution Address:  
广东工业大学计算机学院, 广州, 广东 510006, 中国  
Institution Address:  
School of Computer Science, Guangdong University of Technology, Guangzhou, Guangdong 510006, China  
Institution:  
广东工业大学计算机学院  
Institution:  
School of Computer Science, Guangdong University of Technology

### 8. Intelligent fault diagnosis of rolling bearings based on continuous

摘要: Accurate fault diagnosis is critical to operating rotating machinery safely and efficiently. Traditional fault information description methods rely on experts to extract statistical features, which inevitably leads to the problem of information loss. As a result, this paper proposes an intelligent fault diagnosis of rolling bearings based on a continuous wavelet transform(CWT)-multiscale feature fusion and an improved channel attention mechanism. Different from traditional CNNs, CWT can convert the 1-D signals into 2-D images, and extract the wavelet power spectrum, which is conducive to model recognition. In this case, the multiscale feature fusion was implemented by the parallel 2-D convolutional neural networks to accomplish deeper feature fusion. Meanwhile, the channel attention mechanism is improved by converting from compressed to extended ways in the excitation block to better obtain the evaluation score of the channel. The proposed model has been validated using two bearing datasets, and the results show that it has excellent accuracy compared to existing methods.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
9  
ISSN:  
1507-2711  
入藏号:  
WOS:001061827600005  
Indexed Date:  
2023-09-21  
Institution Address:  
Northeastern Univ, Sch Mech Engn & Automat, Shenyang 110819, Peoples R China  
Institution Address:  
Northeastern Univ, Key Lab Vibrat & Control Aero Prop Syst, Minist Educ, Shenyang 110819, Peoples R China  
Institution Address:  
Northeastern Univ, Key Lab Multidisciplinary Design Optimizat Comple, Shenyang 110819, Peoples R China  
Institution Address:  
Angang Steel Co Ltd, Anshan 114021, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Angang Steel Co Ltd

### 9. A Deeply Supervised Convolutional Neural Network for Pavement Crack

摘要: Automatic crack detection is vital for efficient and economical road maintenance. With the explosive development of convolutional neural networks (CNNs), recent crack detection methods are mostly based on CNNs. In this article, we propose a deeply supervised convolutional neural network for crack detection via a novel multiscale convolutional feature fusion module. Within this multiscale feature fusion module, the high-level features are introduced directly into the low-level features at different convolutional stages. Besides, deep supervision provides integrated direct supervision for convolutional feature fusion, which is helpful to improve model convergency and final performance of crack detection. Multiscale convolutional features learned at different convolution stages are fused together to robustly represent cracks, whose geometric structures are complicated and hardly captured by single-scale features. To demonstrate its superiority and generalizability, we evaluate the proposed network on three public crack data sets, respectively. Sufficient experimental results demonstrate that our method outperforms other state-of-the-art crack detection, edge detection, and image segmentation methods in terms of F1-score and mean IU.  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
94  
在中国科学引文数据库中的被引频次:  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
97  
ISSN:  
2162-237X  
eISSN:  
2162-2388  
入藏号:  
WOS:000732294700001  
PubMed ID:  
33720835  
Indexed Date:  
2021-03-12  
Institution Address:  
Chongqing Univ Posts & Telecommunicat, Coll Comp Sci & Technol, Chongqing 400065, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommunicat, Coll Mobile Telecommunicat, Chongqing 401520, Peoples R China  
Institution:  
Institution:

### 10. MSViT: Training Multiscale Vision Transformers for Image Retrieval

摘要: The recently developed vision transformer (ViT) has achieved promising results on image retrieval compared to convolutional neural networks. However, most of these vision transformer-based image retrieval methods use the original ViT model to extract global features, ignoring the importance of local features for image retrieval. In this work, we propose a vision transformer-based multiscale feature fusion image retrieval method (MSViT) to achieve the fusion of global features with local features. The challenge of this research work is how to learn the feature representation ability of transformer model, so as to improve the performance of image retrieval model. First, a transformer-based two-branch network structure is proposed to obtain different scale features by processing image patches with different granularities. Second, we present a multiscale feature fusion strategy, which can efficiently and effectively fuse the feature information of different sizes on two branches. Finally, to more fully utilize the label information to supervise the network training process, we optimize the construction rules for the triplet data. The comparison of experimental results with ten CNN-based and six transformer-based image retrieval methods on four publicly available image datasets shows that our method outperforms the state-of-the-art methods. And ablation experiments show that the designed multiscale feature fusion strategy and improved triplet loss function have an implicit improvement on the performance of MSViT.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
被引频次合计:  
8  
ISSN:  
1520-9210  
eISSN:  
1941-0077  
入藏号:  
WOS:001173299400005  
Indexed Date:  
2024-04-26  
Institution Address:  
Xinjiang Univ, Sch Informat Sci & Engn, Urumqi 830046, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, Sch Comp & Artificial Intelligence, Chengdu 610031, Peoples R China  
Institution Address:  
Xinjiang Univ, Sch Software, Urumqi 830046, Peoples R China  
Institution:  
Institution:  
Institution:

### 11. Encoder- and Decoder-Based Networks Using Multiscale Feature Fusion and

摘要: With the development of convolutional neural networks, the semantic segmentation of remote sensing images has been widely developed, but there are still some unsolved problems in this field due to the lack of multiscale information and the feature mismatch at the upsampling process. To solve these problems, we propose a network called multiscale feature fusion and alignment network (MFANet). MFANet is composed of an encoder and a decoder. The encoder contains a fully convolutional network, a multilevel feature fusion block (MLFFB), and a multiscale feature pyramid (MSFP). These subnetworks can obtain fine-grained feature maps that are full of multiscale and global features and improve segmentation results at multiple object scales. Moreover, MFANet uses a light convolution subnetwork, called decoder, to upsample the segmentation map stage by stage. Combining three scales of features, the decoder can promote the feature alignment at the upsampling stage. Along with the decoder, MFANet utilizes a multistage supervision loss to enhance the localization performance and boundary regression ability. Benefitting from the encoder and decoder structure and the innovative components inside encoder, MFANet is very powerful for the semantic segmentation of remote sensing images and can suit the complicated environment. We evaluate our MFANet on the Vaihingen and Potsdam data sets, and it outperforms the state-of-art methods both in the metric and visual effect.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
11  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000665034700008  
Indexed Date:  
2021-07-08  
Institution Address:  
Beihang Univ, Inst Elect Informat Engn, Beijing 100191, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Sch Sci, Xian 710049, Peoples R China  
Institution:  
Institution:

### 12. A Lightweight Multiscale Feature Fusion Network for Remote Sensing

摘要: In recent decades, remote sensing object counting has attracted increasing attention from academia and industry due to its potential benefits in urban traffic, public safety, and road planning. However, this issue is becoming a challenge for computer vision because of various technical barriers, such as large-scale variation, complex background interference, and nonuniform density distribution. Recent results show hopeful prospects for object counting using convolutional neural networks (CNNs), but most existing CNN-based methods draw on larger and more complex architectures, which leads to a huge computational and storage burdens, severely limiting their application in real-world scenarios. In this article, a lightweight multiscale feature fusion network for remote sensing object counting, named LMSFFNet, is presented to achieve a better balance between the running speed of the network and the counting accuracy. Specifically, in the encoding process, we select a MobileViT module as the backbone of the network to reduce the numbers of network parameters and computing cost. In return, a cascade structure of the channel-spatial attention mechanisms compensates for the weaker feature extraction ability of the lightweight network. In the decoding process, a lightweight multiscale context fusion module (LMCFM) as a multiscale feature fusion module is developed to solve the problem that the number of parameters increases with the expansion of the object scale when extracting multiscale features. In addition, a lightweight counting scale pooling module (LCSPM) is used to mine the subtle features of the target object. Two kinds of typical object counting experiments, namely, experiments on remote sensing benchmarks (RSOC dataset) and crowd benchmarks (ShanghaiTech, UCF-QNRF, and UCF\_CC\_50 datasets), show the effectiveness of the proposed method.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
19  
被引频次合计:  
20  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000925114100004  
Indexed Date:  
2023-03-08  
Institution Address:  
Chongqing Univ Sci & Technol, Coll Intelligent Technol & Engn, Chongqing 401331, Peoples R China  
Institution Address:  
Chongqing Coll Elect Engn, Inst Big Data & Optimizat, Chongqing 401331, Peoples R China  
Institution:  
Institution:

### 13. FTMF-Net: A Fourier Transform-Multiscale Feature Fusion Network for

摘要: The detection and resection of small polyp objects in colonoscopy images are of great significance for the prevention of colorectal cancer (CRC). At present, blurred edges, variable lesion shapes, and intraclass dissimilarity pose challenges for accurately segmenting small polyp objects. In recent years, many deep learning methods based on convolutional neural networks (CNNs) have been proposed and successfully applied to polyp segmentation tasks. However, these methods still have three limitations: 1) limited ability to mine boundary detail information; 2) insufficient ability to capture rich global context information; and 3) introduced additional complex feature extraction operations. To alleviate these challenges, we propose a Fourier transform-multiscale feature fusion network (FTMF-Net) for segmentation of small polyp objects. The core idea includes two points: 1) Fourier transform (FT) module extracts more detailed boundary information and 2) multiscale feature fusion (MFF) module enriches global semantic feature information. FTMF-Net mainly has the following advantages: 1) the proposed model has excellent performance for small polyp object segmentation; 2) this method greatly reduces the complexity of the model without significantly increasing the number of network parameters; and 3) the network is relatively simple and easy to understand. Extensive experiments with 11 state-of-the-art (SOTA) methods on five small polyp object datasets show that our proposed FTMF-Net has superior segmentation performance.  
在 BIOSIS Citation Index 中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
15  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001036137100020  
Indexed Date:  
2023-08-21  
Institution Address:  
Henan Normal Univ, Coll Comp & Informat Engn, Xinxiang 453000, Peoples R China  
Institution Address:  
Key Lab Artificial Intelligence & Personalized Lea, Xinxiang 453000, Peoples R China  
Institution:  
Institution:  
Key Lab Artificial Intelligence & Personalized Lea

### 14. NAS-MFF: NAS-Guided Multiscale Feature Fusion Network With Pareto

摘要: Underwater target recognition technology based on sonar images has received considerable critical attention in recent years. However, the sonar sensors encounter disturbance from seafloor reverberation noise and a complicated background, resulting in notable difficulties for precise sonar target classification. On the other hand, traditional machine learning methods inevitably lose features relying on expert systems, and manual network creation is relatively inefficient with limited sonar data. To tackle these challenges, we propose a neural architecture search (NAS)-guided multiscale feature fusion (NAS-MFF) algorithm for sonar images classification based on the differentiable architecture search. Specifically, our approach consists of two stages: a search stage with the Pareto optimization, and a training stage using the optimal architecture. NAS-MFF begins by reconfiguring the search space based on the characteristics of sonar images, which includes the introduction of the MF Block k with multiscale feature extraction ability. By synergizing a recognition-driven convolutional neural network (CNN) with Pareto optimization, it achieves a dual advantage in both accuracy and model efficiency using the available data. Extensive experiments on three sonar image datasets of different sizes and sources distinctly demonstrate that NAS-MFF outperforms several existing manual design methods and NAS approaches.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001219652600097  
Indexed Date:  
2024-06-09  
Institution Address:  
Northwestern Polytech Univ, Sch Marine Sci & Technol, Xian 710072, Shaanxi, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Ocean Inst, Taicang 215400, Jiangsu, Peoples R China  
Institution:  
Institution:

### 15. Genetic Feature Fusion for Object Skeleton Detection

摘要: Object skeleton detection requires the convolutional neural networks to recognize objects and their parts in the cluttered background, overcome the image definition degradation brought by the pooling layers, and predict the location of skeleton pixels in different scale granularity. Most existing object skeleton detection methods take great efforts into the designing of side-output networks for multiscale feature fusion. Despite the great progress achieved by them, there are still many problems that hinder the development of object skeleton detection, such as the manually designed network is labor-intensive and the network initialization depends on models pretrained on large-scale datasets. To alleviate these issues, we propose a genetic NAS method to automatically search on a newly designed architecture search space for adaptive multiscale feature fusion. Furthermore, we introduce a symmetric encoder-decoder search space based on reversing the VGG network, in which the decoder can reuse the ImageNet pretrained model of VGG. The searched networks improve the performance of the state-of-the-art methods on commonly used skeleton detection benchmarks, which proves the efficacy of our method.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1939-0114  
eISSN:  
1939-0122  
入藏号:  
WOS:000631865300003  
Indexed Date:  
2021-03-30  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 100049, Peoples R China  
Institution:

### 16. Multiscale Local and Global Feature Fusion for the Detection of Steel

摘要: Steel surface defects have a significant impact on the quality and performance of many industrial products and cause huge economic losses. Therefore, it is meaningful to detect steel surface defects in real time. To improve the detection performance of steel surface defects with variable scales and complex backgrounds, in this paper, a novel method for detecting steel surface defects through a multiscale local and global feature fusion mechanism is proposed. The proposed method uses a convolution operation with a downsampling mechanism in the convolutional neural network model to obtain rough multiscale feature maps. Then, a context-extraction block (CEB) is proposed to adopt self-attention learning on the feature maps extracted by the convolution operation at each scale to obtain multiscale global context information to make up for the shortcomings of convolutional neural networks (CNNs), thus forming a novel multiscale self-attention mechanism. Afterwards, using the feature pyramid structure, multiscale feature maps are fused to improve multiscale object detection. Finally, the channel and spatial attention module and the WIOU (Wise Intersection over Union) loss function are introduced. The model achieved 78.2% and 71.9% mAP respectively on the NEU-DET and GC10-DET dataset. Compared to algorithms such as Faster RCNN and EDDN, this method is effective in improving the detection performance of steel surface defects.  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
7  
eISSN:  
2079-9292  
入藏号:  
WOS:001035129800001  
Indexed Date:  
2023-08-04  
Institution Address:  
Xinyang Normal Univ, Sch Comp & Informat Technol, Xinyang 464000, Peoples R China  
Institution Address:  
Zhengzhou Univ, Sch Comp & Artificial Intelligence, Zhengzhou 450001, Peoples R China  
Institution Address:  
Zhengzhou Natl Supercomp Ctr, Zhengzhou 450001, Peoples R China  
Institution Address:  
Engn Res Ctr Intelligent Cluster Syst, Minist Educ, Zhengzhou 450001, Peoples R China  
Institution:  
Institution:  
Institution:  
Zhengzhou Natl Supercomp Ctr  
Institution:  
Engn Res Ctr Intelligent Cluster Syst

### 17. Malicious Code Variant Identification Based on Multiscale Feature Fusion

摘要: The increasing volume and types of malwares bring a great threat to network security. The malware binary detection with deep convolutional neural networks (CNNs) has been proved to be an effective method. However, the existing malware classification methods based on CNNs are unsatisfactory to this day because of their poor extraction ability, insufficient accuracy of malware classification, and high cost of detection time. To solve these problems, a novel approach, namely, multiscale feature fusion convolutional neural networks (MFFCs), was proposed to achieve an effective classification of malware based on malware visualization utilizing deep learning, which can defend against malware variants and confusing malwares. The approach firstly converts malware code binaries into grayscale images, and then, these images will be normalized in size by utilizing the MFFC model to identify malware families. Comparative experiments were carried out to verify the performance of the proposed method. The results indicate that the MFFC stands out among the recent advanced methods with an accuracy of 98.72% and an average cost of 5.34 milliseconds on the Malimg dataset. Our method can effectively identify malware and detect variants of malware families, which has excellent feature extraction capability and higher accuracy with lower detection time.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1687-5265  
eISSN:  
1687-5273  
入藏号:  
WOS:000739019900009  
PubMed ID:  
34950195  
Indexed Date:  
2022-01-11  
Institution Address:  
Air Force Engn Univ, Coll Air & Missile Def, Xian 710051, Peoples R China  
Institution:

### 18. Three dimensional-CNN Classification Method of Mural Multispectral Image

摘要: The classification and recognition of pigments is the basis of ancient mural protection and restoration. The multispectral imaging method can quickly obtain and analyze the spectral image data of mural pigments without damage. Continuous convolution and pooling operations in the traditional convolutional neural network feature extraction algorithm will lose part of the feature information of the fresco multispectral image, making the image details unable to be reconstructed, resulting in an unsmooth boundary of the classified image. To solve this problem, a three-dimensional hole convolution residual neural network based on multiscale feature fusion is proposed to classify multispectral mural images. To begin, the hole structure is introduced into the convolution kernel to improve the receptive field and extract different scale information to avoid the loss of some features caused by the pooling operation. Second, the feature fusion method is used to combine images of different scales. Finally, a multilevel gradient of the feature map is introduced to prevent the edge from disappearing. On the multispectral image dataset of simulated murals, the experimental results show that the proposed method's overall accuracy and average accuracies are 98.87% and 96.89%, respectively. The proposed method not only outperforms the control groups in classification accuracy, but it also produces classification images with clearer boundaries.  
摘要:  
颜料的分类识别是古代壁画进行保护修复的基础,多光谱成像方法能够无损快速地获取壁画颜料的光谱图像数据并进行分析。传统利用卷积神经网络进行特征提取的算法中连续的卷积和池化操作会丢失壁画多光谱图像的部分特征信息,使得图像细节无法重建,导致分类图像边界不平滑。针对该问题,提出了一种基于多尺度特征融合的三维空洞卷积残差神经网络对壁画多光谱图像进行颜料分类。首先,在卷积核中引入空洞结构提高卷积核的感受野来提取不同尺度信息,避免池化操作所导致的部分特征丢失;其次,使用特征融合的方法融合不同尺度的特征图,增加多尺度特征的结构层次;最后,引入残差学习模块避免网络层数加深导致的梯度消失问题,重建完整的边缘信息。实验结果表明,所提方法在模拟壁画多光谱图像数据集上的总体精度和平均精度分别达到了98.87%和96.89%,与各对照组相比,不仅具有更好的分类精度,而且得到了边界更清晰的分类图像。  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7370125  
Indexed Date:  
2023-04-22  
Institution Address:  
西安建筑科技大学信息与控制工程学院, 西安, 陕西 710055, 中国  
Institution Address:  
陕西省文物保护研究院, 西安, 陕西 710075, 中国  
Institution Address:  
School of Information and Control Engineering,Xi'an University of Architecture and Technology, Xi'an, Shaanxi 710055, China  
Institution Address:  
Shaanxi Provincial Institute of Cultural Relics Protection, Xi'an, Shaanxi 710075, China  
Institution:  
西安建筑科技大学信息与控制工程学院  
Institution:  
陕西省文物保护研究院  
Institution:  
School of Information and Control Engineering,Xi'an University of Architecture and Technology  
Institution:  
Shaanxi Provincial Institute of Cultural Relics Protection

### 19. Remote Sensing Image Scene Classification Using Multiscale Feature

摘要: In remote sensing scene classification (RSSC), features can be extracted with different spatial frequencies where high-frequency features usually represent detailed information and low-frequency features usually represent global structures. However, it is challenging to extract meaningful semantic information for RSSC tasks by just utilizing high- or low-frequency features. The spatial composition of remote sensing images (RSIs) is more complex than that of natural images, and the scales of objects vary significantly. In this article, a multiscale feature fusion covariance network (MF(2)CNet) with octave convolution (Oct Conv) is proposed, which can extract multifrequency and multiscale features from RSIs. First, the multifrequency feature extraction (MFE) module is used to obtain fine-grained frequency features by Oct Conv. Then, the features of different layers in MF(2)CNet are fused by the multiscale feature fusion (MF2) module. Finally, instead of using global average pooling (GAP), global covariance pooling (GCP) extracts high-order information from RSIs to capture richer statistics of deep features. In the proposed MF(2)CNet, the obtained multifrequency and multiscale features can effectively improve the performance of CNNs. Experimental results on four public RSI datasets show that MF(2)CNet has advantages in RSSC over current state-of-the-art methods. The source codes of this method can be found at <uri>https://github.com/liuqingxin-chd/MF2CNet</uri>.  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
43  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
被引频次合计:  
44  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000783579800033  
Indexed Date:  
2022-05-09  
Institution Address:  
Changan Univ, Sch Elect & Control Engn, Xian 710064, Peoples R China  
Institution Address:  
Univ New South Wales, Sch Engn & Informat Technol, Canberra, ACT 2612, Australia  
Institution:  
Institution:

### 20. Multiscale Feature Fusion and Anchor Adaptive Object Detection Algorithm

摘要: Aiming at the problems of low detection accuracy resulting from insufficient feature extraction and inaccurate detection box positioning in the Faster R-CNN algorithm, an object detection algorithm based on multiscale feature fusion and anchor adaptation is proposed. First, the high- and low-level features between adjacent levels were fully extracted using the two-way fusion method; then, the multiscale features were balanced so that the integrated features could obtain the same amount of semantic information and detailed information with different resolutions, improving the object recognition ability. Finally, the anchor was generated by adaptively predicting the position and shape of the anchor using the characteristic information of the object in the region proposals network(RPN). The experimental results of the algorithm based on VOC dataset show that compared with the Faster R-CNN algorithm based on ResNet50, the multiscale feature fusion strategy in the proposed algorithm strengthens the detection ability for objects with different scales. The adaptive anchor mechanism can improve the positioning accuracy and avoid missed detection of small objects, and the overall detection results of the proposed algorithm have good performances. The proposed algorithm improves the average detection accuracy by approximately 3. 20 percentage points.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
1006-4125  
入藏号:  
WOS:000823188100042  
Indexed Date:  
2022-08-10  
Institution Address:  
Anhui Jianzhu Univ, Sch Mech & Elect Engn, Hefei 230601, Anhui, Peoples R China  
Institution Address:  
Key Lab Intelligent Mfg Construct Machinery, Hefei 230601, Anhui, Peoples R China  
Institution Address:  
Anhui Jianzhu Univ, Key Lab Construct Machinery Fault Diag & Early Wa, Hefei 230601, Anhui, Peoples R China  
Institution:  
Institution:  
Key Lab Intelligent Mfg Construct Machinery  
Institution:

### 21. Convolutional Neural Network with Multiscale Fusion and Attention

摘要: Melanoma segmentation based on a convolutional neural network (CNN) has recently attracted extensive attention. However, the features captured by CNN are always local that result in discontinuous feature extraction. To solve this problem, we propose a novel multiscale feature fusion network (MSFA-Net). MSFA-Net can extract feature information at different scales through a multiscale feature fusion structure (MSF) in the network and then calibrate and restore the extracted information to achieve the purpose of melanoma segmentation. Specifically, based on the popular encoder-decoder structure, we designed three functional modules, namely MSF, asymmetric skip connection structure (ASCS), and calibration decoder (Decoder). In addition, a weighted cross-entropy loss and two-stage learning rate optimization strategy are designed to train the network more effectively. Compared qualitatively and quantitatively with the representative neural network methods with encoder-decoder structure, such as U-Net, the proposed method can achieve advanced performance.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
1687-5265  
eISSN:  
1687-5273  
入藏号:  
WOS:000829490400002  
PubMed ID:  
35747726  
Indexed Date:  
2022-08-03  
Institution Address:  
Chongqing Univ Sci & Technol, Sch Intelligent Technol & Engn, Chongqing 401331, Peoples R China  
Institution Address:  
Luzhou Vocat & Tech Coll, Liquor Making Microbial Applicat & Detect Technol, Luzhou 646000, Sichuan, Peoples R China  
Institution:  
Institution:  
Luzhou Vocat & Tech Coll

### 22. AMFFNet: Asymmetric Multiscale Feature Fusion Network of RGB-NIR for

摘要: Multimodal fusion object detection aims to improve detection accuracy by integrating information from multiple modalities. RGB image-based object detection makes finely sorting solid waste according to material challenging. This article built a dual-camera acquisition platform using a line-scan color camera and a hyperspectral camera to collect RGB and hyperspectral images. In order to use RGB images and hyperspectral images for feature fusion more effectively, we propose an asymmetric multiscale feature fusion network (AMFFNet) based on RGB-near-infrared (NIR) multisensor fusion technology. Specifically, we designed a hyperspectral image convolution unit (HICU) to fully extract multiscale features from hyperspectral images. Second, we concatenate the hyperspectral feature maps and the feature maps output by feature pyramid networks (FPNs) of the RGB image feature extraction stage to achieve asymmetric multiscale feature fusion. In addition, a dimensionality reduction strategy (DRS) is proposed to remove the bands that are redundant and have a low signal-to-noise ratio in hyperspectral images. Ablation studies have confirmed the effectiveness of AMFFNet components, and we have conducted extensive integration experiments on the article-based solid waste dataset. The experimental results and analysis show that AMFFNet based on mask region-based convolutional neural network (Mask RCNN), faster RCNN, and RetinaNet detectors outperforms the original model by 5.05%, 3.57%, and 6.36% on AP at IoU = 0.5 and by 3.1%, 2.05%, and 3.73% on COCO's standard AP metric, and the parameters of the network increased by 0.47-0.56 M. In conclusion, the proposed method effectively improves the performance of the object detection model for fine identification of solid waste.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001049981700037  
Indexed Date:  
2023-08-31  
Institution Address:  
Huaqiao Univ, Fujian Prov Univ, Huaqiao Univ, Sch Electromech & Automat,Key Lab Proc Monitoring, Xiamen 361021, Peoples R China  
Institution Address:  
Xiamen Luhai Environm Protect Co Ltd, Xiamen 361001, Peoples R China  
Institution:  
Institution:  
Xiamen Luhai Environm Protect Co Ltd

### 23. OrchidNet: A Self-Supervised Learning-Based Efficient Multiscale Feature

摘要: Classifying orchids is a challenging task: the color, shape, and texture features of each part of an orchid image are small, the category similarity is high, and the features are blurred. Traditional deep learning methods are not suitable for orchid image classification, mainly because they rely on a large amount of human interpretation for labeling models to achieve a certain degree of accuracy. Therefore, to address the human interpretation problem of orchid classification, we propose a self-supervised learning-based efficient multiscale feature fusion convolutional neural lightweight network architecture called OrchidNet. OrchidNet is divided into three parts: 1) a convolutional neural network (CNN) attention network; 2) a self-supervised fine-grained mechanism; and 3) a feature pyramid network for multiscale feature fusion. In this article, we focus on lightweight and high-performance materials. The motivation for our method is to use CNNs to upsample the backbone architecture (backbone) of local visual features. In orchid images, most of the features are extracted by the CNN backbone (attention mechanism). Afterward, the self-supervised fine-grained mechanism is used for feature screening. Finally, the global and local feature information is downsampled through the feature pyramid for multiscale feature fusion. This study significantly impacts the image-processing community of orchid classification. The proposed OrchidNet improves the accuracy of orchid classification and uses a self-supervised learning method that does not require manual labeling, thereby reducing the labeling time and the cost of manual interpretation. The experimental results show that the proposed OrchidNet is superior to state-of-the-art methods for orchid classification used in recent years.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2327-4662  
入藏号:  
WOS:001433294700001  
Indexed Date:  
2025-03-08  
Institution Address:  
Natl Penghu Univ Sci & Technol, Dept Comp Sci & Informat Engn, Penghu 880011, Taiwan  
Institution:  
Natl Penghu Univ Sci & Technol

### 24. A Multiscale Dual-Branch Feature Fusion and Attention Network for

摘要: Recently, hyperspectral image classification based on deep learning has achieved considerable attention. Many convolutional neural network classification methods have emerged and exhibited superior classification performance. However, most methods focus on extracting features by using fixed convolution kernels and layer-wise representation, resulting in feature extraction singleness. Additionally, the feature fusion process is rough and simple. Numerous methods get accustomed to fusing different levels of features by stacking modules hierarchically, which ignore the combination of shallow and deep spectral-spatial features. In order to overcome the preceding issues, a novel multiscale dual-branch feature fusion and attention network is proposed. Specifically, we design a multiscale feature extraction (MSFE) module to extract spatial-spectral features at a granular level and expand the range of receptive fields, thereby enhancing the MSFE ability. Subsequently, we develop a dual-branch feature fusion interactive module that integrates the residual connection's feature reuse property and the dense connection's feature exploration capability, obtaining more discriminative features in both spatial and spectral branches. Additionally, we introduce a novel shuffle attention mechanism that allows for adaptive weighting of spatial and spectral features, further improving classification performance. Experimental results on three benchmark datasets demonstrate that our model outperforms other state-of-the-art methods while incurring the lower computational cost.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
36  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
37  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000692210900007  
Indexed Date:  
2021-09-12  
Institution Address:  
Hohai Univ, Coll Comp & Informat, Nanjing 211100, Peoples R China  
Institution:

### 25. Three dimensional-CNN Classification Method of Mural Multispectral Image

摘要: The classification and recognition of pigments is the basis of ancient mural protection and restoration. The multispectral imaging method can quickly obtain and analyze the spectral image data of mural pigments without damage. Continuous convolution and pooling operations in the traditional convolutional neural network feature extraction algorithm will lose part of the feature information of the fresco multispectral image, making the image details unable to be reconstructed , resulting in an unsmooth boundary of the classified image. To solve this problem , a three-dimensional hole convolution residual neural network based on multiscale feature fusion is proposed to classify multispectral mural images. To begin, the hole structure is introduced into the convolution kernel to improve the receptive field and extract different scale information to avoid the loss of some features caused by the pooling operation. Second , the feature fusion method is used to combine images of different scales. Finally, a multilevel gradient of the feature map is introduced to prevent the edge from disappearing. On the multispectral image dataset of simulated murals, the experimental results show that the proposed method's overall accuracy and average accuracies are 98.87% and 96.89%, respectively. The proposed method not only outperforms the control groups in classification accuracy, but it also produces classification images with clearer boundaries.  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1006-4125  
入藏号:  
WOS:000892702100040  
Indexed Date:  
2022-12-23  
Institution Address:  
Xian Univ Architecture & Technol, Sch Informat & Control Engn, Xian 710055, Shaanxi, Peoples R China  
Institution Address:  
Shaanxi Prov Inst Cultural Rel Protect, Xian 710075, Shaanxi, Peoples R China  
Institution:  
Institution:  
Shaanxi Prov Inst Cultural Rel Protect

### 26. IBA-U-Net: Attentive BConvLSTM U-Net with Redesigned Inception for

摘要: Accurate segmentation of medical images plays an essential role in their analysis and has a wide range of research and application values in fields of practice such as medical research, disease diagnosis, disease analysis, and auxiliary surgery. In recent years, deep convolutional neural networks have been developed that show strong performance in medical image segmentation. However, because of the inherent challenges of medical images, such as irregularities of the dataset and the existence of outliers, segmentation approaches have not demonstrated sufficiently accurate and reliable results for clinical employment. Our method is based on three key ideas: (1) integrating the BConvLSTM block and the Attention block to reduce the semantic gap between the encoder and decoder feature maps to make the two feature maps more homogeneous, (2) factorizing convolutions with a large filter size by Redesigned Inception, which uses a multiscale feature fusion method to significantly increase the effective receptive field, and (3) devising a deep convolutional neural network with multiscale feature fusion and a Attentive BConvLSTM mechanism, which integrates the Attentive BConvLSTM block and the Redesigned Inception block into an encoder-decoder model called Attentive BConvLSTM U-Net with Redesigned Inception (IBA-U-Net). Our proposed architecture, IBA-U-Net, has been compared with the U-Net and state-of-the-art segmentation methods on three publicly available datasets, the lung image segmentation dataset, skin lesion image dataset, and retinal blood vessel image segmentation dataset, each with their unique challenges, and it has improved the prediction performance even with slightly less calculation expense and fewer network parameters. By devising a deep convolutional neural network with a multiscale feature fusion and Attentive BConvLSTM mechanism, medical image segmentation of different tasks can be completed effectively and accurately with only 45% of U-Net parameters.  
在中国科学引文数据库中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
23  
在 BIOSIS Citation Index 中的被引频次:  
9  
被引频次合计:  
26  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:000687883000002  
PubMed ID:  
34157471  
Indexed Date:  
2021-09-01  
Institution Address:  
Nanchang Univ, Sch Informat Engn, Nanchang 330031, Jiangxi, Peoples R China  
Institution Address:  
Carleton Univ, Dept Syst & Comp Engn, Ottawa, ON K1S 5B6, Canada  
Institution:  
Institution:

### 27. PMSNet: Multiscale Partial-Discharge Signal Feature Recognition Model

摘要: Partial discharge (PD) is a localized discharge phenomenon in the insulator of electrical equipment resulting from the electric field strength exceeding the local dielectric breakdown electric field. Partial-discharge signal identification is an important means of assessing the insulation status of electrical equipment and critical to the safe operation of electrical equipment. The identification effect of traditional methods is not ideal because the PD signal collected is subject to strong noise interference. To overcome noise interference, quickly and accurately identify PD signals, and eliminate potential safety hazards, this study proposes a PD signal identification method based on multiscale feature fusion. The method improves identification efficiency through the multiscale feature fusion and feature aggregation of phase-resolved partial-discharge (PRPD) diagrams by using PMSNet. The whole network consists of three parts: a CNN backbone composed of a multiscale feature fusion pyramid, a down-sampling feature enhancement (DSFB) module for each layer of the pyramid to acquire features from different layers, a Transformer encoder module dominated by a spatial interaction-attention mechanism to enhance subspace feature interactions, a final categorized feature recognition method for the PRPD maps and a final classification feature generation module (F-Collect). PMSNet improves recognition accuracy by 10% compared with traditional high-frequency current detection methods and current pulse detection methods. On the PRPD dataset, the validation accuracy of PMSNet is above 80%, the validation loss is about 0.3%, and the training accuracy exceeds 85%. Experimental results show that the use of PMSNet can greatly improve the recognition accuracy and robustness of PD signals and has good practicality and application prospects.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:001245270800001  
PubMed ID:  
38894132  
Indexed Date:  
2024-06-21  
Institution Address:  
Wuhan Text Univ, Sch Elect & Elect Engn, Wuhan 430200, Peoples R China  
Institution Address:  
Wuhan Text Univ, State Key Lab New Text Mat & Adv Proc Technol, Wuhan 430200, Peoples R China  
Institution Address:  
Cent China Normal Univ, Fac Artificial Intelligence Educ, Wuhan 430079, Peoples R China  
Institution:  
Institution:  
Institution:

### 28. Semantic Segmentation of Test Papers Based on Subspace Multi-Scale

摘要: A key step for achieving the semantic segmentation of the test paper is to separate the printed and handwritten regions. In order to improve the effect of the semantic segmentation of the test paper, this paper proposes an improved attention algorithm based on the MaskRCNN network. By embedding the Subspace Multiscale Feature Fusion (SMFF) module into the feature pyramid structure of the MaskRCNN network, the attention features are calculated via the subspace such that the spatial and channel redundancy in the feature map can be reduced. By multiscale feature fusion, the features of text regions with different sizes can be effectively extracted and the correlation between features can be enhanced. The experimental results show that for the target detection and semantic segmentation tasks of the test paper image dataset, the MaskRCNN network model based on the SMFF module can increase the average accuracy by 15.8% and 10.2% higher than that of the original MaskRCNN network model. Moreover, it also has greater performance improvement than the MaskRCNN based on the commonly used attention module.  
摘要:  
分离印刷体和手写体区域是实现试卷语义分割的关键步骤,为了提升试卷语义分割的效果,提出一种基于MaskRCNN网络的注意力改进算法。该算法将子空间多尺度特征融合(Subspace Multiscale Feature Fusion, SMFF)模块嵌入MaskRCNN网络的特征金字塔结构中, SMFF模块基于子空间计算注意力特征,减少特征图中的空间和通道冗余;通过多尺度特征融合,有效提取不同大小文本区域的特征并增强特征间的关联性。实验结果表明,在试卷图像数据集的目标检测和语义分割任务上,基于SMFF模块的MaskRCNN网络模型比MaskRCNN原网络模型的平均准确率分别提高了15.8%和10.2%,比基于常用注意力模块的MaskRCNN网络也有较大的性能提升。  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-3080  
入藏号:  
CSCD:7522355  
Indexed Date:  
2023-11-10  
Institution Address:  
华东理工大学信息科学与工程学院, 上海 200237, 中国  
Institution Address:  
School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China  
Institution:  
华东理工大学信息科学与工程学院  
Institution:  
School of Information Science and Engineering, East China University of Science and Technology

### 29. Semantic Segmentation of LiDAR Point Cloud Based on CAFF-PointNet

摘要: Herein, we propose a convolutional neural network based on channel attention mechanism for multiscale feature fusion regarding the characteristics of LiDAR point clouds, such as the complex geometric structure and extreme scale variations among different categories, resulting in the issue of low classification accuracy of small targets. First, low-level features (planarity, linearity, normal vector, and eigen entropy) are calculated for each point by setting a spherical neighborhood, and they are fused with high-level features acquired by the network to improve the geometry awareness of the constructed model. Then, a multiscale feature fusion module is designed based on the channel attention mechanism to learn fusion weight coefficient so that the network can adapt to the receptive field size of different scale objects and realize different scales information filtering, which improves the classification performance of the small-scale object. According to the experiments, the average F, score using the ISPRS Vaihingen 3D Semantic Labeling benchmark is 72.2 % Compared with other algorithms, our model has the highest classification accuracy in the powerline and car categories with F-1 scores of 64.3 % and 79.9 %, respectively.  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1006-4125  
入藏号:  
WOS:000752877700050  
Indexed Date:  
2022-02-26  
Institution Address:  
Shanghai Jiao Tong Univ, Sch Elect Informat & Elect Engn, Dept Instrument Sci & Engn, Shanghai 200240, Peoples R China  
Institution:

### 30. RSMNet: A Robust Stacked Multiscale Feature Fusion Network for Visible

摘要: The aircrafts and ships in visible remote sensing (RS) images are of different scales. They are difficult to detect as they may be easily obscured by complex weather conditions, such as snow and cloud. Therefore, it is important to eliminate the interference of complex weather conditions in order to detect these multiscale objects accurately. This letter proposes an improved robust stacked multiscale feature fusion network RSMNet to address this problem from two aspects. First, a stacked dilated convolution is used to enlarge the receptive fields of high-resolution images and improve the ability to extract multiscale information. Second, the maps of extracted features are resized and integrated to refine the connection among different layers. Compared to the original faster R-CNN model, RSMNet provides a 2% and 3.9% higher AP in the detection of aircrafts and ships, respectively. RSMNet also shows much more robust performance than the original model in detection under cloudy and snowy conditions.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001080604700007  
Indexed Date:  
2023-10-24  
Institution Address:  
Northeastern Univ Qinhuangdao, Hebei Key Lab Micronano Precis Opt Sensing & Measu, Qinhuangdao 066004, Peoples R China  
Institution Address:  
Northeastern Univ Qinhuangdao, Sch Control Engn, Qinhuangdao 066004, Peoples R China  
Institution Address:  
PLA Army Engn Univ, Shijiazhuang 210007, Peoples R China  
Institution:  
Institution:  
Institution:

### 31. M-FFN: multi-scale feature fusion network for image captioning

摘要: In this work, we present a novel multi-scale feature fusion network (M-FFN) for image captioning task to incorporate discriminative features and scene contextual information of an image. We construct multi-scale feature fusion network by leveraging spatial transformation and multi-scale feature pyramid networks via feature fusion block to enrich spatial and global semantic information. In particular, we take advantage of multi-scale feature pyramid network to incorporate global contextual information by employing atrous convolutions on top layers of convolutional neural network (CNN). And, the spatial transformation network is exploited on early layers of CNN to remove intra-class variability caused by spatial transformations. Further, the feature fusion block integrates both global contextual information and spatial features to encode the visual information of an input image. Moreover, spatial-semantic attention module is incorporated to learn attentive contextual features to guide the captioning module. The efficacy of the proposed model is evaluated on the COCO dataset.  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000801084600001  
Indexed Date:  
2022-06-06  
Institution Address:  
Indian Inst Technol Hyderabad, Hyderabad, India  
Institution:

### 32. Underwater Target Detection Lightweight Algorithm Based on Multi-Scale

摘要: The performance of underwater target detection algorithms is affected by poor imaging quality in underwater environments. Due to the arithmetic power limitation of underwater devices, existing deep learning networks are unable to provide efficient detection processes with high detection accuracy. Lightweight CNN models have been actively applied for underwater environment detection, yet their lite feature fusion networks cannot provide effective fusion effects and reduce the detection accuracy. In this paper, a lightweight algorithm based on multi-scale feature fusion was proposed, with the model parameters greatly reduced, improving the target detection accuracy. The forward propagation memory overhead is reduced by using multi-scale shared convolutional kernels and pooling operations to co-construct the query matrix in the Tansformer encoding stage. Then, the feature fusion path is optimized in order to enhance the connection of multi-scale features. A multiscale feature adaptive fusion strategy is used to enhance the detection performance and reduce the dependence on the complex feature extraction network. The feature extraction network is also reparameterized to simplify the operation. Using the UPRC offshore dataset for validation, the study results have demonstrated that the statistical mAP metrics validate the detection accuracy. Compared with SSD, RetinaNet and YOLOv5-s improved by 13%, 8.6%, and 0.8%, while the number of parameters decreased by 76.09%, 89.74%, and 87.67%. In addition, compared with the YOLOv5-lite model algorithm with the same parameter volume, the mAP is improved by 3.8%, which verifies the accuracy and efficiency of the algorithm in this paper.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
16  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
17  
eISSN:  
2077-1312  
入藏号:  
WOS:000940441300001  
Indexed Date:  
2023-03-18  
Institution Address:  
Hunan Univ Sci & Technol, Sch Informat & Elect Engn, Xiangtan 411201, Peoples R China  
Institution:

### 33. MFCANet: Multiscale Feature Context Aggregation Network for Oriented

摘要: Rotated object detection in remote sensing images presents a highly challenging task due to the extensive fields of view and complex backgrounds. While Convolutional Neural Networks (CNNs) and Transformer networks have made progress in this area, there is still a lack of research on extracting and fusing features for small targets in complex backgrounds. To address this gap, we have extended the RTMDet framework by introducing three modules: the Focused Feature Context Aggregation Module, the Feature Context Information Enhancement Module, and the Multi-scale Feature Fusion Module. In the Focused Feature Context Aggregation Module, we replaced the Spatial Pyramid Pooling Bottleneck (SPPFBottleneck) to better extract small target features by focusing on contextual information. The Feature Context Information Enhancement Module enhances the model's perception of multi-dimensional temporal and spatial information. Finally, we combined the original features with the fused ones to prevent the loss of specific features during the fusion process. Our proposed model, named the Multi-scale Feature Context Aggregation Network (MFCANet), was evaluated on four challenging remote sensing datasets (MAR20, SRSDD, HRSC, and DIOR-R). The experimental results demonstrate that our method outperforms baseline models, achieving improvements of 2.13%, 10.28%, 1.46%, and 1.13% in mAP for the MAR20, SRSDD, HRSC, and DIOR-R datasets, respectively.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
2169-3536  
入藏号:  
WOS:001196446500001  
Indexed Date:  
2024-04-07  
Institution Address:  
Anhui Tech Coll Mech & Elect Engn, Sch Internet & Commun, Wuhu 241003, Peoples R China  
Institution Address:  
State Gride Wuhu Power Supply Co, Wuhu 230061, Peoples R China  
Institution Address:  
Hefei Univ Technol, Sch Instrument Sci & Optoelect Engn, Hefei 230009, Peoples R China  
Institution Address:  
Anhui Tech Coll Mech & Elect Engn, Sch Mech Engn, Wuhu 241003, Peoples R China  
Institution:  
Anhui Tech Coll Mech & Elect Engn  
Institution:  
State Gride Wuhu Power Supply Co  
Institution:  
Institution:  
Anhui Tech Coll Mech & Elect Engn

### 34. LKAFFNet: A Novel Large-Kernel Attention Feature Fusion Network for Land

摘要: The accurate segmentation of land cover in high-resolution remote sensing imagery is crucial for applications such as urban planning, environmental monitoring, and disaster management. However, traditional convolutional neural networks (CNNs) struggle to balance fine-grained local detail with large-scale contextual information. To tackle these challenges, we combine large-kernel convolutions, attention mechanisms, and multi-scale feature fusion to form a novel LKAFFNet framework that introduces the following three key modules: LkResNet, which enhances feature extraction through parameterizable large-kernel convolutions; Large-Kernel Attention Aggregation (LKAA), integrating spatial and channel attention; and Channel Difference Features Shift Fusion (CDFSF), which enables efficient multi-scale feature fusion. Experimental comparisons demonstrate that LKAFFNet outperforms previous models on both the LandCover dataset and WHU Building dataset, particularly in cases with diverse scales. Specifically, it achieved a mIoU of 0.8155 on the LandCover dataset and 0.9326 on the WHU Building dataset. These findings suggest that LKAFFNet significantly improves land cover segmentation performance, offering a more effective tool for remote sensing applications.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001393887300001  
PubMed ID:  
39796845  
Indexed Date:  
2025-01-17  
Institution Address:  
Macao Polytech Univ, Fac Appl Sci, Macau 999078, Peoples R China  
Institution Address:  
Macao Polytech Univ, Macau 999078, Peoples R China  
Institution:  
Institution:

### 35. A ViT-Based Multiscale Feature Fusion Approach for Remote Sensing Image

摘要: Semantic segmentation plays an indispensable role in automatic analysis of remote sensing image data. However, the abundant semantic information and irregular shape patterns in remote sensing images are difficult to utilize, making it hard to segment remote sensing images only using convolution and single-scale feature maps. To achieve better segmentation performance, a multiscale feature pyramid decoder (MFPD) is proposed to fuse image features extracted by vision transformer (ViT). The decoder employs a novel 2-D-to-3-D transform method to obtain multiscale feature maps that contain rich context information and fuses the multiscale feature maps by channel concatenation. Furthermore, a dimension attention module (DAM) is designed to further aggregate the context information of the extracted remote sensing image features. This approach yields superior mean intersection over union (mIoU) on the Gaofen2-CZ dataset (60.42%) and GID-5 dataset (68.21%). Experimental results indicate that the comprehensive performance of our approach exceeds the compared segmentation methods based on convolutional neural network (CNN) and ViT.  
Web of Science 核心合集中的 "被引频次":  
28  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
被引频次合计:  
28  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000838380900007  
Indexed Date:  
2022-08-21  
Institution Address:  
Changsha Univ Sci & Technol, Coll Comp & Commun, Changsha 410114, Peoples R China  
Institution:

### 36. MS3D: A Multi-Scale Feature Fusion 3D Object Detection Method for

摘要: With advancements in autonomous driving, LiDAR has become central to 3D object detection due to its precision and interference resistance. However, challenges such as point cloud sparsity and unstructured data persist. This study introduces MS3D (Multi-Scale Feature Fusion 3D Object Detection Method), a novel approach to 3D object detection that leverages the architecture of a 2D Convolutional Neural Network (CNN) as its core framework. It integrates a Second Feature Pyramid Network to enhance multi-scale feature representation and contextual integration. The Adam optimizer is employed for efficient adaptive parameter tuning, significantly improving detection performance. On the KITTI dataset, MS3D achieves average precisions of 93.58%, 90.91%, and 88.46% in easy, moderate, and hard scenarios, respectively, surpassing state-of-the-art models like VoxelNet, SECOND, and PointPillars.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001366817000001  
Indexed Date:  
2024-12-07  
Institution Address:  
Jimei Univ, Chengyi Coll, Xiamen 361021, Peoples R China  
Institution Address:  
Jimei Univ, Sch Ocean Informat Engn, Xiamen 361021, Peoples R China  
Institution:  
Institution:

### 37. A semantic segmentation algorithm using multi-scale feature fusion with

摘要: The advancement of deep learning has boosted the research on image semantic segmentation.At present,most effective methods for this research are based on the fully convolutional neural networks.Although the existing semantic segmentation methods can effectively segment the image as a whole,they cannot clearly identify the edge information of the overlapped objects in the image,and cannot effectively fuse the high- and low-layer feature information of the image.To address the above problems,superpixel segmentation was employed as an auxiliary optimization to optimize the segmentation results of object edges based on the fully convolutional neural network.At the same time,the design of a joint cross-stage partial multiscale feature fusion module can enable the utilization of image spatial information.In addition,a skip structure was added to the upsampling module to enhance the learning ability of the network,and two loss functions were adopted to ensure network convergence and improve network performance.The network was trained and tested on the public datasets PASCAL VOC 2012.Compared with other image semantic segmentation methods,the proposed network can improve the accuracies in pixel and segmentation,and displays strong robustness.  
摘要:  
深度学习的发展加快了图像语义分割的研究。目前,最有效的图像语义分割研究方法大部分都是基于全卷积神经网络(FCNN),尽管现有的语义分割方法能有效地对图像进行整体分割,但对于图像中的重叠遮挡物体不能清晰地识别出边缘信息,也不能有效地融合图像高低层的特征信息。针对以上问题,在采用FCNN来解决图像语义分割问题的基础上,利用超像素分割对物体边缘的特殊优势作为辅助优化,对粗糙分割结果进行优化。同时在FCNN中利用空洞卷积设计了一个联合局部跨阶段的多尺度特征融合模块,其能有效地利用图像的空间信息。此外还在网络的上采样模块中加入跳跃连接结构,用来增强网络的学习能力,在训练过程中采用2个损失函数来保证网络稳定收敛和提升网络的性能,图像语义分割网络在公开的数据集PASCAL VOC 2012上进行训练测试。实验结果表明,该改进算法在像素精度和分割准确率方面均有提升,且具有较强的鲁棒性。  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
5  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
2095-302X  
入藏号:  
CSCD:7000268  
Indexed Date:  
2021-09-25  
Institution Address:  
上海师范大学信息与机电工程学院, 上海 200234, 中国  
Institution Address:  
上海交通大学电子信息与电气工程学院, 上海 200240, 中国  
Institution Address:  
College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai 200234, China  
Institution Address:  
College of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China  
Institution:  
上海师范大学信息与机电工程学院  
Institution:  
上海交通大学电子信息与电气工程学院  
Institution:  
College of Information, Mechanical and Electrical Engineering, Shanghai Normal University  
Institution:  
College of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University

### 38. A multiscale feature fusion method for cursive text detection in natural

摘要: Text detection in natural images is a challenging problem due to variations in text size, aspect ratio, alignment and background complexity. This paper proposes a multiscale feature fusion convolutional neural network method to detect cursive and multi-language text in natural images. The proposed method combines VGG-16 features at multi-scales and multi-layers and creates a new convolutional feature map of shallow and deep layers. On top of convolutional feature map, a vertical text proposal generation method is used that generates fixed-size text proposals. A recurrent layer is implemented which takes the convolutional feature maps of 3 x 3 window as sequential input and updates the recurrent state internally in the hidden layers. The output of recurrent layer is mapped to the two fully connected layers to predict the text/non-text region proposals and bounding boxes regression. The model is evaluated on a custom-developed Urdu scene text dataset and the ICDAR-MLT17 Arabic text image dataset.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1368-2199  
eISSN:  
1743-131X  
入藏号:  
WOS:000909910100001  
Indexed Date:  
2023-01-26  
Institution Address:  
Univ New South Wales, Sch Engn & Informat Technol, Canberra, Australia  
Institution Address:  
Quaid E Awam Univ, Dept Informat Technol, Nawabshah 67450, Pakistan  
Institution Address:  
Quaid E Awam Univ, Dept Comp Syst Engn, Nawabshah, Pakistan  
Institution Address:  
BUPT, State Key Lab Networking & Switching Technol, Beijing, Peoples R China  
Institution:  
Institution:  
Quaid E Awam Univ  
Institution:  
Quaid E Awam Univ  
Institution:

### 39. LMFNet: A Lightweight Multiscale Fusion Network With Hierarchical

摘要: Three-dimensional (3-D) face recognition (FR) can improve the usability and user-friendliness of human-machine interaction. In general, 3-D FR can be divided into high-quality and low-quality 3-D FR according to different interaction scenarios. The low-quality data can be easily obtained, so its application prospect is more extensive. However, the challenge is how to balance the trade-offs between data accuracy and real-time performance. To solve this problem, we propose a lightweight multiscale fusion network (LMFNet) with a hierarchical structure based on single-mode data for low-quality 3-D FR. First, we design a backbone network with only five feature extraction blocks to reduce computational complexity and improve the inference speed. Second, we devise a mid-low adjacent layer with a multiscale feature fusion (ML-MSFF) module to extract the facial texture and contour information, and a mid-high adjacent layer with a multiscale feature fusion (MH-MSFF) module to obtain the discriminative information in high-level features. Then, a hierarchical multiscale feature fusion (HMSFF) module is formed by combining these two modules mentioned above to acquire the local information of different scales. Finally, we enhance the expression of features by integrating HMSFF with a global convolutional neural network for improving recognition accuracy. Experiments on Lock3DFace, KinectFaceDB, and IIIT-D datasets demonstrate that our proposed LMFNet can achieve superior performance on low-quality datasets. Furthermore, experiments on the cross-quality database based on Bosphorus and the different intensity noise low-quality datasets based on UMB-DB and Bosphorus show that our network is robust and has a high generalization ability. It satisfies the real-time requirement, which lays a foundation for a smooth and user-friendly interactive experience.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
11  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
2168-2291  
eISSN:  
2168-2305  
入藏号:  
WOS:000852228900001  
Indexed Date:  
2022-09-17  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Elect Engn, Beijing Key Lab Work Safety Intelligent Monitorin, Beijing 100876, Peoples R China  
Institution Address:  
Univ Portsmouth, Sch Creat Technol, Portsmouth PO1 2DJ, Hants, England  
Institution:  
Institution:

### 40. A deep recursive multi-scale feature fusion network for image

摘要: Recently, Convolutional Neural Networks (CNNs) have achieved great success in Single Image Super-Resolution (SISR). In particular, the recursive networks are now widely used. However, existing recursion-based SISR networks can only make use of multi-scale features in a layer-wise manner. In this paper, a Deep Recursive Multi-Scale Feature Fusion Network (DRMSFFN) is proposed to address this issue. Specifically, we propose a Recursive Multi-Scale Feature Fusion Block (RMSFFB) to make full use of multi-scale features. Besides, a Progressive Feature Fusion (PFF) technique is proposed to take advantage of the hierarchical features from the RMSFFB in a global manner. At the reconstruction stage, we use a deconvolutional layer to upscale the feature maps to the desired size. Extensive experimental results on benchmark datasets demonstrate the superiority of the proposed DRMSFFN in comparison with the state-of-the-art methods in both quantitative and qualitative evaluations.  
Web of Science 核心合集中的 "被引频次":  
14  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
16  
ISSN:  
1047-3203  
eISSN:  
1095-9076  
入藏号:  
WOS:000906300300001  
Indexed Date:  
2023-01-15  
Institution Address:  
Sichuan Univ, Coll Elect & Informat Engn, Chengdu 610065, Peoples R China  
Institution Address:  
Univ Ghent, Dept Data Anal & Math Modelling, KERMIT, Coupure Links 653, B-9000 Ghent, Belgium  
Institution:  
Institution:

### 41. A Multibranch Network With Multilayer Feature Fusion for No-Reference

摘要: With the widespread application of digital images in various domains, the accurate measurement of image quality has become particularly crucial. This article introduces a novel multibranch multilayer feature fusion network (MFFNet) to address the inadequate expression of multiscale and semantic features and local visual feature consideration in existing no-reference image quality assessment (NR-IQA) algorithms. MFFNet comprises a primary and a sub-branch. Through convolutional neural network (CNN) feature extraction, the main branch uses a multiscale feature enhancement (MSFE) module to capture fine-grained features at each layer, thus significantly enhancing its capability to represent local features. It subsequently merges these distinct-scale features through the multilayer feature fusion (MLFF) module to improve MFFNet performance. Recognizing human attention to the local image area during image quality evaluation, the sub-branch acquires local visual information using a classical superpixel segmentation model. Finally, the two branches are fused using an element-by-element multiplication operation. Comparative experiments are conducted using four representative datasets-CSIQ, TID2013, LIVEC, and CID2013-demonstrating that the MFFNet method outperforms most advanced techniques, thereby establishing the method's effectiveness.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
4  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001239831400020  
Indexed Date:  
2024-06-10  
Institution Address:  
North China Elect Power Univ, Sch Control & Comp Engn, Baoding 071000, Hebei, Peoples R China  
Institution Address:  
North China Elect Power Univ, Hebei Key Lab Knowledge Comp Energy & Power, Baoding 071000, Hebei, Peoples R China  
Institution Address:  
North China Elect Power Univ, Sch Elect & Elect Engn, Baoding 071000, Hebei, Peoples R China  
Institution:  
Institution:  
Institution:

### 42. Multiscale feature fusion network for automatic port segmentation from

摘要: In recent years, remote sensing image observation technology has developed rapidly. Extracting coastlines from remote sensing images has become an indispensable means of port area measurement. Port segmentation from remote sensing images is an important method of coastline extraction and measurement. The remote sensing images used are panchromatic remote sensing images. Due to complex remote sensing image scenes and the large difference in feature information at different scales, traditional segmentation methods cannot perform effective extraction, and it is difficult to accurately segment the coastline in remote sensing images. We propose a multiscale feature fusion network for automatic port segmentation from remote sensing images. First, to reduce the redundant parameters and complex operation problems in traditional convolutional neural networks, we propose using MobileNetv2 as the base network for feature extraction to achieve a lightweight model. Then aiming at the feature differences of remote sensing images at different scales, we present atrous convolution as a convolution method for the entire network and combine a multiscale feature fusion method to extract the features of remote sensing images and improve the feature extraction ability. To reduce the problem of ships calling at the port being easily mistaken for port area and causing false segmentation, we propose a method of eliminating the ship area to reduce the interference. Finally, the comprehensive evaluation of a large number of Google port remote sensing data shows that compared with existing methods, the proposed method has the characteristics of being lightweight and having high precision. (c) 2022 Society of Photo-Optical Instrumentation Engineers (SPIE)  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
1931-3195  
入藏号:  
WOS:000917345800008  
Indexed Date:  
2023-02-17  
Institution Address:  
North China Univ Technol, Sch Informat Sci & Technol, Beijing, Peoples R China  
Institution Address:  
China Acad Space Technol, Gen Dept Remote Sensing Satellites, Beijing, Peoples R China  
Institution:  
Institution:  
China Acad Space Technol

### 43. Hierarchical Shrinkage Multiscale Network for Hyperspectral Image

摘要: Recently, deep learning (DL)-based hyperspectral image classification (HSIC) has attracted substantial attention. Many works based on the convolutional neural network (CNN) model have been certificated to be significantly successful for boosting the performance of HSIC. However, most of these methods extract features by using a fixed convolutional kernel and ignore multiscale features of the ground objects of hyperspectral images (HSIs). Although some recent methods have proposed multiscale feature extraction schemes, more computing and storage resources were consumed. Moreover, when using CNN to implement HSI classification, many methods only use the high-level semantic information extracted from the end of the network, ignoring the edge information extracted from shallow layers of the network. To settle the preceding two issues, a novel HSIC method based on hierarchical shrinkage multiscale network and the hierarchical feature fusion is proposed, with which the newly proposed classification framework can fuse features generated by both of multiscale receptive field and multiple levels. Specifically, multidepth and multiscale residual block (MDMSRB) is constructed by superposition dilated convolution to realize multiscale feature extraction. Furthermore, according to the change of feature size in different stages of the neural networks, we design a hierarchical shrinkage multiscale feature extraction network by pruning MDMSRB to reduce the redundancy of network structure. In addition, to make full use of the features extracted in each stage of the network, the proposed network hierarchically integrates low-level edge features and high-level semantic features effectively. Experimental results demonstrate that the proposed method achieves more competitive performance with a limited computational cost than other state-of-the-art methods.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
13  
被引频次合计:  
15  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000663535500011  
Indexed Date:  
2021-07-09  
Institution Address:  
Hohai Univ, Coll Comp & Informat, Nanjing 211100, Peoples R China  
Institution:

### 44. IX-ResNet: fragmented multi-scale feature fusion for image

摘要: With the continuous in-depth study of convolutional neural network in computer vision, how to improve the performance of network structure has been the focus of current research. Recent works have shown that multi-scale feature concatenation, shortcut connection and grouping convolution can effectively train deeper networks and improve the accuracy and effectiveness of the network. In this paper, we present a novel feature transformation strategy of fragmented multi-scale feature fusion. Moreover, an efficient modularized image classification network, IX-ResNet, is proposed based on this new strategy. IX-ResNet consists of many large isomorphic modules stacked in the form of residual network while Each large module can be composed of many small heterogeneous modules. The performance of IX-ResNet is verified on cifar-10, cifar-100 and ImageNet-1 K datasets, which indicates that IX-ResNet model using fragmented multi-scale feature fusion strategy can further improve accuracy compare to the original grouping convolution network ResNeXt with the same or even lower parameters.  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:000655178800002  
Indexed Date:  
2021-06-01  
Institution Address:  
Xian Polytech Univ, Xian, Peoples R China  
Institution:

### 45. Multiscale Feature Fusion for Skin Lesion Classification

摘要: Skin cancer has a high mortality rate, and early detection can greatly reduce patient mortality. Convolutional neural network (CNN) has been widely applied in the field of computer-aided diagnosis. To improve the ability of convolutional neural networks to accurately classify skin lesions, we propose a multiscale feature fusion model for skin lesion classification. We use a two-stream network, which are a densely connected network (DenseNet-121) and improved visual geometry group network (VGG-16). In the feature fusion module, we construct multireceptive fields to obtain multiscale pathological information and use generalized mean pooling (GeM pooling) to reduce the spatial dimensionality of lesion features. Finally, we built and tested a system with the developed skin lesion classification model. The experiments were performed on the dataset ISIC2018, which can achieve a good classification performance with a test accuracy of 91.24% and macroaverages of 95%.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
10  
ISSN:  
2314-6133  
eISSN:  
2314-6141  
入藏号:  
WOS:000913693600001  
PubMed ID:  
36644161  
Indexed Date:  
2023-02-01  
Institution Address:  
Anhui Jianzhu Univ, Coll Elect & Informat Engn, Hefei 230000, Peoples R China  
Institution Address:  
Anhui Jianzhu Univ, Anhui Int Joint Res Ctr Ancient Architecture Intel, Hefei 230000, Peoples R China  
Institution Address:  
Anhui Jianzhu Univ, Anhui Prov Key Lab Intelligent Bldg & Bldg Energy, Hefei 230000, Peoples R China  
Institution:  
Institution:  
Institution:

### 46. A High-Stability Diagnosis Model Based on a Multiscale Feature Fusion

摘要: Recently, the diagnosis of rotating machines based on deep learning models has achieved great success. Many of these intelligent diagnosis models are assumed that training and test data are subject to independent identical distributions (IIDs). Unfortunately, such an assumption is generally invalid in practical applications due to noise disturbances and changes in workload. To address the above problem, this article presents a high-stability diagnosis model named the multiscale feature fusion convolutional neural network (MFF-CNN). MFF-CNN does not rely on tedious data preprocessing and target domain information. It is composed of multiscale dilated convolution, self-adaptive weighting, and the new form of maxout (NFM) activation. It extracts, modulates, and fuses the input samples' multiscale features so that the model focuses more on the health state difference rather than the noise disturbance and workload difference. Two diagnostic cases, including noisy cases and variable load cases, are used to verify the effectiveness of the present model. The results show that the present model has a strong health state identification capability and anti-interference capability for variable loads and noise disturbances.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
19  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
19  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000693606300013  
Indexed Date:  
2021-09-18  
Institution Address:  
Beijing Univ Chem Technol, Sch Mech Elect Engn, Beijing 100029, Peoples R China  
Institution Address:  
Beijing Univ Technol, Key Lab Adv Mfg Technol, Beijing 100124, Peoples R China  
Institution:  
Institution:

### 47. Multiscale holospectrum convolutional neural network-based fault

摘要: In recent years, the convolutional neural network (CNN) has been widely used in the field of intelligent fault diagnosis. However, each convolutional layer of CNN cannot take the overall and local information into account, and the feature extraction ability of CNN with fewer layers is weak. These circumstances lead to poor performance of CNN in practical fault diagnosis with variable operating conditions. To solve these problems, this paper proposes a multiscale holospectrum CNN (MH-CNN) based on the methods of two-dimensional multiscale feature fusion and decision-level feature fusion. First, the continuous wavelet transform is used to map the time-domain signal to the time-frequency plane to fully reflect the complex information contained in the signal. Then the two-dimensional multiscale feature fusion is introduced to extract features at different scales, which can take both overall and local information into account. Finally the decision-level feature fusion is introduced to fuse the features from signal in X, Y directions in the decision-level of CNN, which serves to enhance the features. By combining these methods, the proposed MH-CNN can extract more distinguishable features with a shallow structure, which can ensure the classification capability while avoiding the overfitting problem caused by overly complex networks. The effectiveness of the MH-CNN is verified using complicated data sets consisting of 16 rolling bearings with four different health conditions, two speeds and three loads. Results show that the proposed MH-CNN achieves a correct rate of 99.8% for rolling bearing fault diagnosis under variable operating conditions, which is much higher than other comparative methods.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
26  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
27  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:000671443400001  
Indexed Date:  
2021-07-16  
Institution Address:  
Xi An Jiao Tong Univ, Sch Mech Engn, State Key Lab Mfg Syst Engn, Xian 710049, Peoples R China  
Institution:

### 48. Multi-Scale Feature Fusion Detection and Recognition Algorithm for

摘要: Foreign objects on the airport runway pose a fatal threat to the take-off and landing of the aircraft, and the way of human inspection of foreign objects is costly and inefficient. It is necessary to use deep learning detection and measurement methods for the detection of foreign objects on the airport runway. Aiming at the problems of missed detection and inaccurate positioning of foreign objects with different sizes and difficult feature extraction, a foreign object detection and recognition algorithm based on multi-scale feature fusion is proposed. Taking multi-scale feature extraction and fusion as the starting point, it proposes multi-point support for spatial attention, strengthens the feature extraction of foreign objects, and can pay enough attention to the abnormalities of different sizes. Using the BiFPN network for feature fusion can fully integrate feature information of different sizes. The experimental results show that the average accuracy reaches 94.7%, which is 5.9 percentage points higher than YOLOv5, and also surpasses YOLOv6, YOLOv7 and Faster R-CNN, and verifies that this algorithm has good application value in the field of foreign object detection on airport runways.  
摘要:  
机场跑道异物对飞机起飞降落存在着致命威胁,而人工巡检异物的方式成本高,效率低,因此将深度学习检测算法用于机场跑道异物检测是必要的。针对异物尺寸大小不一、特征提取难度大而导致的漏检与定位不准问题,提出一种基于多尺度特征融合的机场跑道异物检测与识别算法。以多尺度特征提取与融合为切入点,提出多分支空间注意力,加强对异物的特征提取,同时能够关注不同尺寸大小的异物。采用BiFPN网络进行特征融合,能够高效融合不同尺度的特征信息。实验结果表明,改进后算法平均精度达到94.7%,相比于YOLOv5提高5.9个百分点,也超越了YOLOv6、YOLOv7以及Faster R-CNN,从而验证了该算法在机场跑道异物检测领域有较好的应用价值。  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7804340  
Indexed Date:  
2024-11-23  
Institution Address:  
中国民航大学工程技术训练中心, 天津 300300, 中国  
Institution Address:  
中国民航大学电子信息与自动化学院, 天津 300300, 中国  
Institution Address:  
Engineering Technology Training Center,Civil Aviation University of China, Tianjin 300300, China  
Institution Address:  
College of Electronic Information and Automation,Civil Aviation University of China, Tianjin 300300, China  
Institution:  
中国民航大学工程技术训练中心  
Institution:  
中国民航大学电子信息与自动化学院  
Institution:  
Engineering Technology Training Center,Civil Aviation University of China  
Institution:  
College of Electronic Information and Automation,Civil Aviation University of China

### 49. Road Extraction from GF-1 Remote Sensing Images Based on Dilated

摘要: This paper aimed to solve the problems of road edge detail information loss and inaccurate road extraction due to multiple downsampling operations of the fully convolutional neural network. Thus, a road extraction method of GF-1 remote sensing images based on dilated convolution residual network with multiscale feature fusion is proposed. First, numerous labels for road extraction are generated through visual interpretation. Second, dilated convolution and multiscale feature perception modules are introduced in each residual block of the residual network, namely, ResNet-101, to enlarge the receptive field of the feature points without reducing the feature map resolution and losing the detailed edge information. Third, through superposition fusion and upsampling operations, the road feature maps of various sizes are fused to obtain the feature maps of the original resolution size. Finally, for classification, the feature maps are input into the Sigmoid classifier. The experimental results indicate that the proposed method is more accurate than the conventional fully convolutional neural network models, with the accuracy rate being more than 98%. The proposed method effectively preserves the integrity and detailed edge information of the road area.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1006-4125  
入藏号:  
WOS:000686476300038  
Indexed Date:  
2021-09-01  
Institution Address:  
Liaoning Tech Univ, Sch Geomat, Fuxin 123000, Liaoning, Peoples R China  
Institution Address:  
Minist Nat Resources, Land & Resources Remote Sensing Applicat Ctr, Beijing 100048, Peoples R China  
Institution:  
Institution:

### 50. GLE-net: global-local information enhancement for semantic segmentation

摘要: Remote sensing (RS) images contain a wealth of information with expansive potential for applications in image segmentation. However, Convolutional Neural Networks (CNN) face challenges in fully harnessing the global contextual information. Leveraging the formidable capabilities of global information modeling with Swin-Transformer, a novel RS images segmentation model with CNN (GLE-Net) was introduced. This integration gives rise to a revamped encoder structure. The subbranch initiates the process by extracting features at varying scales within the RS images using the Multiscale Feature Fusion Module (MFM), acquiring rich semantic information, discerning localized finer features, and adeptly handling occlusions. Subsequently, Feature Compression Module (FCM) is introduced in main branch to downsize the feature map, effectively reducing information loss while preserving finer details, enhancing segmentation accuracy for smaller targets. Finally, we integrate local features and global features through Spatial Information Enhancement Module (SIEM) for comprehensive feature modeling, augmenting the segmentation capabilities of model. We performed experiments on public datasets provided by ISPRS, yielding notably remarkable experimental outcomes. This underscores the substantial potential of our model in the realm of RS image segmentation within the context of scientific research.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2045-2322  
入藏号:  
WOS:001342770900067  
PubMed ID:  
39455717  
Indexed Date:  
2024-11-08  
Institution Address:  
Chongqing Univ Sci & Technol, Dept Intelligent Technol & Engn, 20 East Univ Town Rd, Chongqing 401331, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Sch Mech Engn, 28 Xianning West Rd, Xian 710049, Shaanxi, Peoples R China  
Institution Address:  
Luzhou Vocat & Tech Coll, 2 Changqiao Rd, Luzhou 646299, Sichuan, Peoples R China  
Institution:  
Institution:  
Institution:  
Luzhou Vocat & Tech Coll

### 51. Multiscale Feature Fusion Network Incorporating 3D Self-Attention for

摘要: In recent years, the deep learning-based hyperspectral image (HSI) classification method has achieved great success, and the convolutional neural network (CNN) method has achieved good classification performance in the HSI classification task. However, the convolutional operation only works with local neighborhoods, and is effective in extracting local features. It is difficult to capture interactive features over long distances, which affects the accuracy of classification to some extent. At the same time, the data from HSI have the characteristics of three-dimensionality, redundancy, and noise. To solve these problems, we propose a 3D self-attention multiscale feature fusion network (3DSA-MFN) that integrates 3D multi-head self-attention. 3DSA-MFN first uses different sized convolution kernels to extract multiscale features, samples the different granularities of the feature map, and effectively fuses the spatial and spectral features of the feature map. Then, we propose an improved 3D multi-head self-attention mechanism that provides local feature details for the self-attention branch, and fully exploits the context of the input matrix. To verify the performance of the proposed method, we compare it with six current methods on three public datasets. The experimental results show that the proposed 3DSA-MFN achieves competitive classification and highlights the HSI classification task.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
20  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
被引频次合计:  
20  
eISSN:  
2072-4292  
入藏号:  
WOS:000754631400001  
Indexed Date:  
2022-03-02  
Institution Address:  
North Univ China, Sch Instrument & Elect, Taiyuan 030051, Peoples R China  
Institution Address:  
Henan Inst Engn, Sch Elect Informat Engn, Zhengzhou 451191, Peoples R China  
Institution:  
Institution:  
Henan Inst Engn

### 52. A Review on Multiscale-Deep-Learning Applications

摘要: In general, most of the existing convolutional neural network (CNN)-based deep-learning models suffer from spatial-information loss and inadequate feature-representation issues. This is due to their inability to capture multiscale-context information and the exclusion of semantic information throughout the pooling operations. In the early layers of a CNN, the network encodes simple semantic representations, such as edges and corners, while, in the latter part of the CNN, the network encodes more complex semantic features, such as complex geometric shapes. Theoretically, it is better for a CNN to extract features from different levels of semantic representation because tasks such as classification and segmentation work better when both simple and complex feature maps are utilized. Hence, it is also crucial to embed multiscale capability throughout the network so that the various scales of the features can be optimally captured to represent the intended task. Multiscale representation enables the network to fuse low-level and high-level features from a restricted receptive field to enhance the deep-model performance. The main novelty of this review is the comprehensive novel taxonomy of multiscale-deep-learning methods, which includes details of several architectures and their strengths that have been implemented in the existing works. Predominantly, multiscale approaches in deep-learning networks can be classed into two categories: multiscale feature learning and multiscale feature fusion. Multiscale feature learning refers to the method of deriving feature maps by examining kernels over several sizes to collect a larger range of relevant features and predict the input images' spatial mapping. Multiscale feature fusion uses features with different resolutions to find patterns over short and long distances, without a deep network. Additionally, several examples of the techniques are also discussed according to their applications in satellite imagery, medical imaging, agriculture, and industrial and manufacturing systems.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
45  
在 BIOSIS Citation Index 中的被引频次:  
6  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
45  
eISSN:  
1424-8220  
入藏号:  
WOS:000867177100001  
PubMed ID:  
36236483  
Indexed Date:  
2022-10-21  
Institution Address:  
Univ Kebangsaan Malaysia, Fac Engn & Built Environm, Dept Elect Elect & Syst Engn, Bangi 43600, Selangor, Malaysia  
Institution Address:  
Univ Syiah Kuala, Fac Engn, Dept Elect & Comp Engn, Kopelma Darussalam 23111, Indonesia  
Institution:  
Institution:

### 53. Hyperspectral Image Classification Based on Multibranch Adaptive Feature

摘要: Convolutional neural networks (CNNs) are widely used in hyperspectral image classification (HSIC) due to their exceptional performance. However, current methods for multiscale feature extraction typically rely on single-branch CNNs, potentially causing interference among features of varying scales. To mitigate this issue, we present a multibranch adaptive feature fusion network (MBAFFN) classification method. MBAFFN enhances feature uniqueness and improves the accuracy and reliability of classification results by extracting information at multiple scales through three parallel branches. Furthermore, to address the challenge of capturing global features within CNNs, we introduce a global detail attention (GDA) mechanism aimed at bolstering the network's capability to capture comprehensive information. In addition, we mitigate the issue of neglecting center-pixel importance in convolution operations through a distance suppression attention (DSA) design. To effectively integrate outcomes from multiple branches, we propose a pixel-based adaptive feature fusion strategy, thereby increasing the proportion of features conducive to improved classification results. Lastly, auxiliary loss functions are employed to train the multibranch network. Experimental results on four benchmark datasets demonstrate the superiority of our approach over several state-of-the-art methods, particularly in managing imbalanced small samples. Furthermore, ablation studies validate the effectiveness of the proposed modules.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001308252000024  
Indexed Date:  
2024-09-18  
Institution Address:  
China Univ Geosci, Sch Geophys & Geomat, Wuhan 430074, Peoples R China  
Institution Address:  
State Key Lab Resources & Environm Informat Syst, Beijing 100101, Peoples R China  
Institution:  
Institution:  
State Key Lab Resources & Environm Informat Syst

### 54. Stacked BNAS: Rethinking Broad Convolutional Neural Network for Neural

摘要: Different from other deep scalable architecture-based NAS approaches, Broad Neural Architecture Search (BNAS) proposes a broad scalable architecture which consists of convolution and enhancement blocks, dubbed Broad Convolutional Neural Network (BCNN), as the search space for amazing efficiency improvement. BCNN reuses the topologies of cells in the convolution block so that BNAS can employ few cells for efficient search. Moreover, multi-scale feature fusion and knowledge embedding are proposed to improve the performance of BCNN with shallow topology. However, BNAS suffers some drawbacks: 1) insufficient representation diversity for feature fusion and enhancement and 2) time consumption of knowledge embedding design by human experts. This paper proposes Stacked BNAS, whose search space is a developed broad scalable architecture named Stacked BCNN, with better performance than BNAS. On the one hand, Stacked BCNN treats mini BCNN as a basic block to preserve comprehensive representation and deliver powerful feature extraction ability. For multi-scale feature enhancement, each mini BCNN feeds the outputs of deep and broad cells to the enhancement cell. For multi-scale feature fusion, each mini BCNN feeds the outputs of deep, broad and enhancement cells to the output node. On the other hand, Knowledge Embedding Search (KES) is proposed to learn appropriate knowledge embeddings in a differentiable way. Moreover, the basic unit of KES is an over-parameterized knowledge embedding module that consists of all possible candidate knowledge embeddings. Experimental results show that 1) Stacked BNAS obtains better performance than BNAS-v2 on both CIFAR-10 and ImageNet, 2) the proposed KES algorithm contributes to reducing the parameters of the learned architecture with satisfactory performance, and 3) Stacked BNAS delivers a state-of-the-art efficiency of 0.02 GPU days.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
2168-2216  
eISSN:  
2168-2232  
入藏号:  
WOS:001006677600001  
Indexed Date:  
2023-06-26  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Multimodal Artificial Intelligence S, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Artificial Intelligence, Beijing 100049, Peoples R China  
Institution Address:  
South China Univ Technol, Sch Comp Sci & Engn, Guangzhou 510641, Peoples R China  
Institution Address:  
Pazhou Lab, Ctr Affect Comp & Gen Models, Guangzhou 510335, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 55. Sparse-View Projection Spectral CT Reconstruction via HAMEN

摘要: Spectral computed tomography (CT) can provide attenuation information at different energy levels, which is essential for material decomposition and tissue discrimination. Sparse-view projection can effectively reduce radiation dose but can cause severe artifacts and noise in the reconstructed spectral CT images. Although deep learning reconstruction methods based on convolutional neural networks can improve the image quality, a loss in the tissue detail features is observed. Therefore, a spectral CT reconstruction method based on a hybrid attention module combined with a multiscale feature fusion edge enhancement network (HAMEN) is proposed. The network first extracts edge features of the input images through the edge enhancement module and concatenates them on the images, enriching the input image information. Next, a hybrid attention module is used to generate channel attention and spatial attention maps, which are used to refine the input features. The multiscale feature fusion mechanism is developed at the encoder, and some skip connections are added to minimize feature loss caused by the stacking of convolutional layers. The experimental results show that the peak signal-tonoise ratio of the CT images obtained using the proposed method is 37.64 dB, and the similarity structural index measure is 0.9935. This method can suppress artifacts and noise caused by sparse-view projection while preserving the tissue detail information. Furthermore, the CT image quality is improved for subsequent diagnosis and other works.  
摘要:  
能谱式计算机断层扫描(CT)可以提供不同能量下的衰减信息,有利于物质分解和组织分辨。稀疏投影可以有效降低辐射剂量,但会导致重建后的CT图像出现严重的伪影和噪声。基于卷积神经网络的深度学习重建方法虽在一定程度上改善了图像质量,却存在组织细节特征丢失严重等问题。提出一个基于混合注意力与多尺度特征融合相结合的边缘增强型网络(HAMEN)的能谱CT重建方法。首先利用边缘增强模块提取边缘特征并将其叠加到图像上,用于扩充输入图像信息;然后采用混合注意力模块分别生成通道注意力图和空间注意力图,以细化输入特征;并在网络的编码器处引入多尺度特征融合机制,增加跳跃连接以减少卷积层堆叠导致的特征丢失。实验结果表明,所提重建方法得到的CT图像的峰值信噪比可达37.64 dB,结构相似性指数达0.9935。此方法可在抑制稀疏投影导致的伪影和噪声的同时最大程度地保留组织细节信息,为后续的诊断等工作提供高质量图像。  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7783044  
Indexed Date:  
2024-10-18  
Institution Address:  
天津大学微电子学院, 天津 300072, 中国  
Institution Address:  
天津大学微电子学院;;天津市成像与感知微电子技术重点实验室, ;;天津市成像与感知微电子技术重点实验室, ;;, 天津;;天津 300072;;300072, 中国  
Institution Address:  
School of Microelectronics,Tianjin University, Tianjin 300072, China  
Institution Address:  
School of Microelectronics,Tianjin University;;Tianjin Key Laboratory of Imaging and Sensing Microelectronic Technology, ;;Tianjin Key Laboratory of Imaging and Sensing Microelectronic Technology, ;;, Tianjin;;Tianjin 300072;;300072  
Institution:  
天津大学微电子学院  
Institution:  
天津大学微电子学院;;天津市成像与感知微电子技术重点实验室  
Institution:  
School of Microelectronics,Tianjin University  
Institution:  
School of Microelectronics,Tianjin University;;Tianjin Key Laboratory of Imaging and Sensing Microelectronic Technology

### 56. Sparse-View Projection Spectral CT Reconstruction via HAMEN

摘要: Spectral computed tomography (CT) can provide attenuation information at different energy levels, which is essential for material decomposition and tissue discrimination. Sparse-view projection can effectively reduce radiation dose but can cause severe artifacts and noise in the reconstructed spectral CT images. Although deep learning reconstruction methods based on convolutional neural networks can improve the image quality, a loss in the tissue detail features is observed. Therefore, a spectral CT reconstruction method based on a hybrid attention module combined with a multiscale feature fusion edge enhancement network (HAMEN) is proposed. The network first extracts edge features of the input images through the edge enhancement module and concatenates them on the images, enriching the input image information. Next, a hybrid attention module is used to generate channel attention and spatial attention maps, which are used to refine the input features. The multiscale feature fusion mechanism is developed at the encoder, and some skip connections are added to minimize feature loss caused by the stacking of convolutional layers. The experimental results show that the peak signal-to-noise ratio of the CT images obtained using the proposed method is 37. 64 dB, and the similarity structural index measure is 0. 9935. This method can suppress artifacts and noise caused by sparse-view projection while preserving the tissue detail information. Furthermore, the CT image quality is improved for subsequent diagnosis and other works.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001284198700023  
Indexed Date:  
2024-09-18  
Institution Address:  
Tianjin Univ, Sch Microelect, Tianjin 300072, Peoples R China  
Institution Address:  
Tianjin Key Lab Imaging & Sensing Microelect Tech, Tianjin 300072, Peoples R China  
Institution:  
Institution:  
Tianjin Key Lab Imaging & Sensing Microelect Tech

### 57. ELCNN: A Deep Neural Network for Small Object Defect Detection of

摘要: Automatic detection of magnetic tile surface defects is an important and challenging task, especially for small objects such as cracks. To address this problem, we propose a novel framework called feature enhancement and loop-shaped fusion convolutional neural network (ELCNN), which enhances shallow features and fuses features with a loop-shaped feature pyramid structure. First, our network focuses on shallow features and maximizes their extraction and enhancement. We believe that if the shallow features cannot be completely extracted and effectively propagated, any further improvements will be futile. Second, inspired by the mechanism of looking and thinking twice, we propose a loop-shaped feature pyramid network (LFPN), which not only applies top-down and bottom-up multiscale feature fusion, but also connects features recursively to the left and right, creatively constructing a loop-shaped feature pyramid structure. The structure enables features to be propagated and fused in all directions, making each layer more directly connected. Just like the shortcut in ResNet, the loop structure is like a loop shortcut in the feature pyramid, which can make feature fusion more efficient and comprehensive. Also, the loop-shortcut can improve the flow of information and gradients throughout the network, making them easy to train. The proposed ELCNN is validated on the production line, and the results demonstrate that our network can significantly improve detection accuracy and meet real-time requirements. We also built a new large-scale dataset magnetic tile-small object dataset (MT-SOD) for the detection of small defects in magnetic tiles and proved its effectiveness in practical applications.  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
23  
被引频次合计:  
23  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000833791900017  
Indexed Date:  
2022-08-10  
Institution Address:  
Tongji Univ, Dept Comp Sci & Technol, Coll Elect & Informat Engn, Shanghai 201804, Peoples R China  
Institution:

### 58. Short-term power prediction of distributed PV based on multi-scale

摘要: To address the challenge of insufficient comprehensive extraction and fusion of meteorological conditions, temporal features, and power periodic features in short-term power prediction for distributed photovoltaic (PV) farms, a TPE-CBiGRU-SCA model based on multiscale feature fusion is proposed. First, multiscale feature fusion of meteorological features, temporal features, and hidden periodic features is performed in PV power to construct the model input features. Second, the relationships between PV power and its influencing factors are modelled from spatial and temporal scales using CNN and Bi-GRU, respectively. The spatiotemporal features are then weighted and fused using the SCA attention mechanism. Finally, TPE-based hyperparameter optimization is used to refine network parameters, achieving PV power prediction for a single field station. Validation with data from a PV field station shows that this method significantly enhances feature extraction comprehensiveness through multiscale fusion at both data and model layers. This improvement leads to a reduction in MAE and RMSE by 26.03% and 38.15%, respectively, and an increase in R2 to 96.22%, representing a 3.26% improvement over other models.  
This paper proposes a TPE-CBiGRU model for short-term power prediction of distributed photovoltaic (PV) systems, addressing challenges in feature extraction and fusion. The model integrates multi-scale features using CNN and Bi-GRU, optimized through Bayesian hyperparameter tuning with TPE. Validation on real PV station data shows significant improvements in prediction accuracy, reducing MAE and RMSE by 26.11% and 35.64% respectively, and increasing R-squared by 3.07% compared to other models. image  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1751-8687  
eISSN:  
1751-8695  
入藏号:  
WOS:001311978700001  
Indexed Date:  
2024-09-21  
Institution Address:  
China Three Gorges Univ, Coll Elect Engn & New Energy, Yichang, Peoples R China  
Institution Address:  
China Three Gorges Univ, Hubei Prov Key Lab Operat & Control, Cascaded Hydropower Stn, Yichang, Peoples R China  
Institution Address:  
Qinghai Univ, Coll Energy & Elect Engn, Xining, Qinghai, Peoples R China  
Institution Address:  
Wuhan Univ, Coll Elect & Automat, Wuhan, Hubei, Peoples R China  
Institution Address:  
State Grid Jibei Elect Power Co, Zhangjiakou Power Supply Co, Zhangjiakou, Hebei, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 59. Visual Object Tracking Based on Mutual Learning Between Cohort

摘要: The deep convolutional neural network (CNN) based tracking-by-detection framework recently has become one of the most popular trackers. However, these methods are either time consuming or have greatly reduced performance. This article aims to achieve nearly identical tracking accuracy with the state-of-the-art CNN tracking-by-detection algorithm with relatively faster speed. We study the existing excellent trackers under the CNN tracking-by-detection framework and introduce the following: a multiscale feature pyramid fusion neural network based on dilated convolutions is constructed to learn a scale-invariant discriminative representation for tracking small objects, a hard-threshold weighted cross-entropy loss function is proposed to decrease the gap between object classification and tracking, and a mutual learning-based training policy is used to fuse the information from the network trained by image patches with different contextual regions to further improve the tracking performance. We conduct comprehensive experiments on visual object tracking benchmarks that validate the achievement of competitive performance of the proposed tracker with relatively faster speed both on qualitative and quantitative criteria. Additionally, the experimental results reveal that the proposed mutual learning-based training policy can accelerate the convergence speed and achieve better generalization performance.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
12  
被引频次合计:  
12  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:000626532100017  
Indexed Date:  
2021-04-16  
Institution Address:  
Xi An Jiao Tong Univ, Sch Elect & Informat Engn, Xian 710049, Peoples R China  
Institution:

### 60. Lightweight Multilevel Feature Fusion Network for Hyperspectral Image

摘要: Hyperspectral images (HSIs), acquired as a 3D data set, contain spectral and spatial information that is important for ground-object recognition. A 3D convolutional neural network (3DCNN) could therefore be more suitable than a 2D one for extracting multiscale neighborhood information in the spectral and spatial domains simultaneously, if it is not restrained by mass parameters and computation cost. In this paper, we propose a novel lightweight multilevel feature fusion network (LMFN) that can achieve satisfactory HSI classification with fewer parameters and a lower computational burden. The LMFN decouples spectral-spatial feature extraction into two modules: point-wise 3D convolution to learn correlations between adjacent bands with no spatial perception, and depth-wise convolution to obtain local texture features while the spectral receptive field remains unchanged. Then, a target-guided fusion mechanism (TFM) is introduced to achieve multilevel spectral-spatial feature fusion between the two modules. More specifically, multiscale spectral features are endowed with spatial long-range dependency, which is quantified by central target pixel-guided similarity measurement. Subsequently, the results obtained from shallow to deep layers are added, respectively, to the spatial modules, in an orderly manner. The TFM block can enhance adjacent spectral correction and focus on pixels that actively boost the target classification accuracy, while performing multiscale feature fusion. Experimental results across three benchmark HSI data sets indicate that our proposed LMFN has competitive advantages, in terms of both classification accuracy and lightweight deep network architecture engineering. More importantly, compared to state-of-the-art methods, the LMFN presents better robustness and generalization.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
12  
被引频次合计:  
12  
eISSN:  
2072-4292  
入藏号:  
WOS:000751090400001  
Indexed Date:  
2022-02-10  
Institution Address:  
Jiangxi Univ Sci & Technol, Sch Informat Engn, Ganzhou 341000, Peoples R China  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Telecommun & Informat Engn, Xian 710121, Peoples R China  
Institution Address:  
Xidian Univ, Sch Artificial Intelligence, Xian 710071, Peoples R China  
Institution:  
Institution:  
Institution:

### 61. LGRF-Net: A Novel Hybrid Attention Network for Lightweight Global Road

摘要: In scenarios where road obstacles complicate feature extraction, designing a lightweight convolutional neural network (CNN) model with minimal parameters and flops while maintaining competitive segmentation accuracy poses one of the most challenging research tasks in remote sensing imaging. Finding the optimal balance between segmentation performance and computational efficiency is crucial. We introduce a novel method for global road feature extraction by strategically employing the light ghost basic block to develop a tiny-ghost link network (TG-LinkNet). A multiscale feature fusion (MSFF) module, which combines the parallel channel position attention mechanism (PCPAM) to deliver accurate road structure information, further supports the goal. We present a solution to the issue of feature fusion information retrieval-induced excessive redundant noise, which might cause serious interference. Furthermore, to efficiently extract edge features and capture long-distance reliance on global features, we create a global context feature extraction (GCFE) module, ultimately resulting in the lightweight global road feature extraction network (LGRF-Net). To facilitate efficient training, we implement a 1:2 weight design within our deep supervision technique, termed hybrid loss (weighted cross entropy (WCE)-Dice). Extensive experiments were conducted on the DeepGlobe (1024 x 1024, 512 x 512) and SpaceNet road datasets. This demonstrates that our network possesses smaller parameters and flops compared to other road-based semantic segmentation methods.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001373843000003  
Indexed Date:  
2024-12-18  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Automat, Xian Key Lab Adv Control & Intelligent Proc, Xian 710061, Peoples R China  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Automat, Xian Robot Intelligent Syst Int Sci & Technol Coop, Xian 710061, Peoples R China  
Institution Address:  
Univ Louisiana Lafayette, Dept Elect & Comp Engn, Lafayette, LA 71465 USA  
Institution:  
Institution:  
Institution:

### 62. Improving the Performance of Convolutional Neural Networks by Fusing

摘要: The width of convolutional neural networks (CNNs) is crucial for improving performance. Many wide CNNs use a convolutional layer to fuse multiscale features or fuse the preceding features to subsequent features. However, these CNNs rarely use blocks, which consist of a series of successive convolutional layers, to fuse multiscale features. In this paper, we propose an approach for improving performance by fusing the low-level features extracted from different blocks. We utilize five different convolutions, including 3 x 3, 5 x 5, 7 x 7,5 x 3 boolean OR 3 x 5 and 7 x 3 boolean OR 3 x 7 , to generate five low-level features, and we design two fusion strategies: low-level feature fusion (L-Fusion) and high-level feature fusion (H-Fusion). Experimental results show that the L-Fusion is more helpful for improving the performance of CNNs, and the 5 x 5 convolution is more suitable for multiscale feature fusion. We summarize the conclusion as a strategy that fuses multiscale features in the preceding stage of CNNs. Furthermore, we propose a new architecture to perceive the input of CNNs by using two self-governed blocks based on the strategy. Finally, we modify five off-the-shelf networks, DenseNet-BC (depth = 40), ALL-CNN-C (depth = 9), Darknet 19 (depth = 19), Resnet 18 (depth = 18) and Resnet 50 (depth = 50), by utilizing the proposed architecture to verify the conclusion, and these updated networks provide more competitive results.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
2169-3536  
入藏号:  
WOS:000652047500001  
Indexed Date:  
2021-05-24  
Institution Address:  
Sichuan Univ, Sch Mech Engn, Chengdu 610065, Peoples R China  
Institution:

### 63. Face Attribute Recognition Combining Feature Fusion and Task Grouping

摘要: To address the issues of insufficient feature extraction and inadequate consideration of attribute correlation in existing face attribute recognition models,a face attribute recognition model combining feature fusion and task grouping is proposed to achieve the simultaneous recognition of multiple face attributes. In the parameter sharing part, features of different scales are fused using multiscale feature fusion modules to enhance feature relevance. Simultaneously,an attribute grouping strategy based on Centered Kernel Alignment-Spectral Clustering(CKA-SC) is designed for attribute identification. The extent of correlation of the attributes is measured using CKA. Based on this,the SC algorithm is utilized to obtain reasonable attribute groups,maximizing the relevance of attributes within the same group to improve attribute recognition accuracy. In the branch part,attention mechanisms are employed to enhance the focus on target regions,and the Uncertainty Weighting(UW) method is used to represent the relative difficulty between tasks. This automatically adjusts the relative weights between the losses of each task group to further optimize the model performance. Experimental results on the publicly available CelebA dataset demonstrate that the proposed model achieves classification accuracy improvements of 0.78,0.09,and 0.02 percentage points compared with the MOON, GNAS,and DMM-CNN models,respectively,with parameter counts accounting for only 1.10%,17.08%,and 0.37% of the mentioned comparison models.  
摘要:  
针对现有人脸属性识别模型存在的特征提取不足、划分属性组时未充分考虑属性相关性强弱等问题,为了实现对多个人脸属性的同时识别,建立一种结合特征融合和任务分组的人脸属性识别模型。在参数共享部分,通过多尺度特征融合模块将不同尺度的特征进行融合增强特征相关性,同时设计基于中心核对齐和谱聚类的属性分组策略进行属性识别,通过中心核对齐方法度量属性的相关程度,并以此为基础使用谱聚类算法得到属性的合理分组,使同一组内的属性相关性尽可能大,提高属性识别准确率。在分支部分,使用注意力机制加强对目标区域的关注,并通过不确定性加权方法表示任务间的相对难度,自动调整每组任务损失之间的相对权重,进一步优化模型性能。在CelebA公开数据集上的实验结果表明,所提模型的分类准确率相较于MOON、GNAS和DMM-CNN模型提升了0.78、0.09和0.02个百分点,参数量仅为上述对比模型的1.10%、17.08%和0.37%。  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1000-3428  
入藏号:  
CSCD:7593405  
Indexed Date:  
2024-01-19  
Institution Address:  
兰州交通大学电子与信息工程学院, 兰州, 甘肃 730070, 中国  
Institution Address:  
兰州交通大学电子与信息工程学院;;甘肃省人工智能与图形图像处理工程研究中心, ;;甘肃省人工智能与图形图像处理工程研究中心, 兰州;;兰州, ;; 730070;;730070  
Institution Address:  
School of Electronic and Information Engineering,Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China  
Institution Address:  
School of Electronic and Information Engineering,Lanzhou Jiaotong University;;Gansu Provincial Engineering Research Center for Artificial Intelligence and Graphic and Image Processing, ;;Gansu Provincial Engineering Research Center for Artificial Intelligence and Graphic and Image Processing, Lanzhou;;Lanzhou, ;; 730070;;730070  
Institution:  
兰州交通大学电子与信息工程学院  
Institution:  
兰州交通大学电子与信息工程学院;;甘肃省人工智能与图形图像处理工程研究中心  
Institution:  
School of Electronic and Information Engineering,Lanzhou Jiaotong University  
Institution:  
School of Electronic and Information Engineering,Lanzhou Jiaotong University;;Gansu Provincial Engineering Research Center for Artificial Intelligence and Graphic and Image Processing

### 64. Change Detection in SAR Images Based on Improved Non-Subsampled Shearlet

摘要: Traditional methods for change detection in synthetic aperture radar images have difficulty in obtaining results from the generated differential image (DI) owing to speckle noise. In recent years, many deep learning-based methods have emerged because of their outstanding anti-noise and self-learning ability. However, they are limited by the requirement of abundant high-precision labels. Therefore, in this article, we propose a novel unsupervised method based on improved non-subsampled shearlet transform (NSST) and multi-scale feature fusion convolutional neural network for change detection. First, this method improves the traditional NSST algorithm and proposes a novel pseudo-label generator to obtain more pseudo-labels with higher confidence. It is noteworthy that the more accurate the pseudo-labels are, the better the change detection results will be. Second, this method designs a multi-scale feature fusion block in the network to make the feature images contain more complete information and reduces the number of pooling layers to avoid losing feature image details. The main idea of this method is to eliminate the step of generating the DI and directly obtain results from the original images. The theoretical analysis and final results conducted on three real datasets prove its validity. Furthermore, to verify the generality and potential of the proposed method, we apply it to the cross-region change detection and compare it with the supervised method, which achieve satisfactory results.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
被引频次合计:  
9  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000728159900018  
Indexed Date:  
2021-12-14  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 101408, Peoples R China  
Institution:  
Institution:

### 65. Multiple disease detection method for greenhouse-cultivated strawberry

摘要: Disease has a significant impact on strawberry quality and yield, and deep learning has become an important approach for the detection of crop disease. To address the problems of complex backgrounds and small disease spots in strawberry disease images from natural environments, we propose a new Faster R\_CNN architecture. The multiscale feature fusion network is composed of ResNet, FPN, and CBAM blocks, and it can effectively extract rich strawberry disease features. We built a dataset for strawberry leaves, flowers and fruits, and the experi-mental results showed that the model was able to effectively detect healthy strawberries and seven strawberry diseases under natural conditions, with an mAP of 92.18% and an average detection time of only 229 ms. The model is compared with Mask R\_CNN and YOLO-v3, and we find that our model can guarantee high accuracy and fast detection operational requirements. Our method provides an effective solution for crop disease detection and can improve farmers' management of the strawberry growing process.  
Web of Science 核心合集中的 "被引频次":  
50  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
4  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
21  
被引频次合计:  
54  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:000826724400001  
Indexed Date:  
2022-07-23  
Institution Address:  
Jiangsu Univ, Key Lab Modern Agr Equipment & Technol, Zhenjiang, Peoples R China  
Institution Address:  
Changzhou Engn & Technol Inst Jiangsu Univ, Changzhou, Peoples R China  
Institution:  
Institution:  
Changzhou Engn & Technol Inst Jiangsu Univ

### 66. Local-Global Multiscale Fusion Network for Semantic Segmentation of

摘要: The extraction of buildings from synthetic aperture radar (SAR) images poses a challenging task in the realm of remote sensing (RS). In recent years, convolutional neural networks (CNNs) have rapidly advanced and found application in the field of RS. Researchers have investigated the potential of CNNs for the semantic segmentation of SAR images, bringing excellent improvements. However, the semantic segmentation of buildings in SAR images still encounters challenges due to the high similarity between features of ground objects and buildings in SAR images, as well as the variability in building structures. In this article, we propose the local-global multiscale fusion network (LGMFNet), based on a dual encoder-decoder structure, for the semantic segmentation of buildings in SAR images. The proposed LGMFNet introduces an auxiliary encoder with a transformer structure to address the limitation of using the main encoder with a CNN structure for global modeling. To embed global dependencies hierarchically into the CNN, we designed the global-local semantic aggregation module (GLSM). The GLSM serves as a bridge between the dual encoders to achieve semantic guidance and coupling from the local to the global level. Furthermore, to bridge the semantic gap between different scales, we designed the multiscale feature fusion network (MSFN) as the decoder. MSFN achieves the interactive fusion of semantic information between various scales by constructing the multiscale feature fusion module. Experimental results demonstrate that the proposed LGMFNet achieves the mIoU of 91.17% on the BIGSARDATA 2023 AISAR competition dataset, outperforming the second-best method by a margin of 0.78%. This evidences the superiority of LGMFNet in comparison to other state-of-the-art methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001200168400009  
Indexed Date:  
2024-04-24  
Institution Address:  
Changshu Inst Technol, Sch Comp Sci & Engn, Suzhou 215500, Peoples R China  
Institution Address:  
China Meteorol Adm, Natl Satellite Meteorol Ctr, Key Lab Radiometr Calibrat & Validat Environm Sate, Beijing 100081, Peoples R China  
Institution Address:  
Innovat Ctr FengYun Meteorol Satellite, Beijing 100081, Peoples R China  
Institution Address:  
Geovis Environm Technol Co Ltd, Technol Res Ctr, Beijing 100094, Peoples R China  
Institution:  
Institution:  
Institution:  
Innovat Ctr FengYun Meteorol Satellite  
Institution:  
Geovis Environm Technol Co Ltd

### 67. Object Detection in Remote Sensing Images Based on Adaptive Multi-Scale

摘要: Multi-scale object detection is critical for analyzing remote sensing images. Traditional feature pyramid networks, which are aimed at accommodating objects of varying sizes through multi-level feature extraction, face significant challenges due to the diverse scale variations present in remote sensing images. This situation often forces single-level features to span a broad spectrum of object sizes, complicating accurate localization and classification. To tackle these challenges, this paper proposes an innovative algorithm that incorporates an adaptive multi-scale feature enhancement and fusion module (ASEM), which enhances remote sensing image object detection through sophisticated multi-scale feature fusion. Our method begins by employing a feature pyramid to gather coarse multi-scale features. Subsequently, it integrates a fine-grained feature extraction module at each level, utilizing atrous convolutions with varied dilation rates to refine multi-scale features, which markedly improves the information capture from widely varied object scales. Furthermore, an adaptive enhancement module is applied to the features of each level by employing an attention mechanism for feature fusion. This strategy concentrates on the features of critical scale, which significantly enhance the effectiveness of capturing essential feature information. Compared with the baseline method, namely, Rotated FasterRCNN, our method achieved an mAP of 74.21% ( 0.81%) on the DOTA-v1.0 dataset and an mAP of 84.90% (+9.2%) on the HRSC2016 dataset. These results validated the effectiveness and practicality of our method and demonstrated its significant application value in multi-scale remote sensing object detection tasks.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
14  
eISSN:  
2072-4292  
入藏号:  
WOS:001183100100001  
Indexed Date:  
2024-03-19  
Institution Address:  
Beijing Univ Posts & Telecommun, Pattern Recognit & Intelligent Vis PRIV, Beijing 100876, Peoples R China  
Institution:

### 68. Multiscale Reference-Aided Attentive Feature Aggregation for Person

摘要: In person re-identification (Re-ID), increasing the diversity of pedestrian features can improve recognition accuracy. In standard convolutional neural networks (CNNs), the receptive fields of neurons in each layer are designed to have the same size. Therefore, in complex pedestrian re-identification tasks, the standard CNNs extract local features but are unable to obtain satisfactory results for global features extracted from the images. Local feature learning methods are helpful for obtaining more abundant features, which focus on the most significant local features and ignore the correlations between features of various parts of the human body. To solve the above problems, a new multiscale reference-aided attentive feature aggregation (MS-RAFA) mechanism is proposed, consisting of three main modules. First, to extract the most significant local features and strengthen the correlations between the features of various parts of the human body, an autoselect module (ASM) is designed, an attentional mechanism that can stack the structural information and spatial relations to form new features. Then, to realize multiscale feature fusion of the multiple output branches of the backbone network and increase feature diversity, we propose a multilayer feature fusion module (MFFM), which enables the model to mine the features hidden by salient features and to learn features better. Finally, to supervise the MFFM and make the network obtain better recognition features, we propose a multiple supervision mechanism. Finally, experimental results demonstrate that our proposed method outperforms the state-of-the-art methods on three large-scale datasets.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:000711717900001  
Indexed Date:  
2021-11-08  
Institution Address:  
Nanchang Hangkong Univ, Key Lab Jiangxi Prov Image Proc & Pattern Recogni, Nanchang 330063, Jiangxi, Peoples R China  
Institution Address:  
Nanchang Hangkong Univ, Sch Software, Nanchang 330063, Jiangxi, Peoples R China  
Institution:  
Institution:

### 69. Underwater Image Enhancement Based on Generate Adversarial Network with

摘要: Aiming at the problems of detail blur, low contrast and blue-green skew in underwater degraded images, an underwater image enhancement algorithm is proposed based on multi-scale feature fusion to generate antagonistic network. Based on the generated adversarial network, this algorithm combines the traditional white balance algorithm and multi-scale enhancement network to enhance underwater degraded images. Firstly, an white balance algorithm with improved channel compensation is used to correct the color bias of underwater degraded images, and the features of the corrected images are extracted by convolutional neural network. Then, the multi-scale features of the image are extracted, and the local features of each layer are enhanced to capture the global features of semantic information by combining with the proposed residual dense blocks, which are fused with the features of the color correction images. Finally, the fusion features are reconstructed into clear images by the reconstruction module, and the details of the images are recovered. The experimental results show that the enhanced underwater image with this algorithm has better defogging effect and more real color, and effectively improves the problem of color bias and blur of underwater image. The experimental results are superior to the comparison algorithm in both subjective and objective indicators.  
摘要:  
针对水下退化图像细节模糊、对比度低和蓝绿色偏问题,提出了一种基于多尺度特征融合生成对抗网络的水下图像增强算法。算法以生成对抗网络为基本框架,结合传统白平衡算法和多尺度增强网络实现对水下退化图像的增强。通过改进的通道补偿白平衡算法矫正蓝绿色偏,并以卷积神经网络提取偏色校正后图像的特征;提取图像多尺度特征,结合提出的残差密集块将每一层的局部特征增强为捕获语义信息的全局特征,并与偏色校正图像的特征相融合;通过重建模块将融合特征重建为清晰图像,恢复图像的细节信息。实验结果表明,该算法增强的水下图像去雾效果较好且颜色更真实,有效改善了水下图像色偏和模糊的问题,在主观指标和客观指标上的实验结果均优于对比算法。  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1002-8331  
入藏号:  
CSCD:7594660  
Indexed Date:  
2024-01-19  
Institution Address:  
安徽理工大学计算机科学与工程学院, 淮南, 安徽 232001, 中国  
Institution Address:  
School of Computer Science and Engineering,Anhui University of Science & Technology, Huainan, Anhui 232001, China  
Institution:  
安徽理工大学计算机科学与工程学院  
Institution:  
School of Computer Science and Engineering,Anhui University of Science & Technology

### 70. An improved deep residual network with multiscale feature fusion for

摘要: Intelligent mechanical fault diagnosis algorithms based on deep learning have achieved considerable success in recent years. However, degradation of the diagnostic accuracy and operational speed has been significant due to unfavorable working conditions and increasing network depth. An improved version of ResNets is proposed in this paper to address these issues. The advantages of the proposed network are presented as follows. Firstly, a multi-scale feature fusion block was designed, to extract multi-scale fault feature information. Secondly, an improved residual block based on depthwise separable convolution was used to improve the operational speed and alleviate the computational burden of the network. The effectiveness of the proposed network was validated by discriminating between diverse health states in a gearbox under normal and noisy conditions. The experimental results show that the proposed network model has a higher classification accuracy than the classical convolutional neural networks, LeNet-5, AlexNet and ResNets and a faster calculation speed than the classical deep neural networks. Furthermore, a visual study of the different stages of the network model was conducted, to effectively comprehend the operational processes of the proposed model.  
在中国科学引文数据库中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
33  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
36  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:000595757300001  
Indexed Date:  
2020-12-18  
Institution Address:  
Shijiazhuang Tiedao Univ, State Key Lab Mech Behav & Syst Safety Traff Engn, Shijiazhuang 050043, Hebei, Peoples R China  
Institution Address:  
Shijiazhuang Tiedao Univ, Sch Mech Engn, Shijiazhuang 050043, Hebei, Peoples R China  
Institution Address:  
Yanshan Univ, Coll Mech Engn, Hebei Prov Key Lab Mech Power & Transmiss Control, Qinhuangdao 066004, Hebei, Peoples R China  
Institution:  
Institution:  
Institution:

### 71. A multiscale feature fusion enhanced CNN with the multiscale channel

摘要: Deep learning (DL) models have been widely used for remote sensing-based landslide mapping due to their impressive capabilities for automatic information extraction. However, the large volumes of parameters and calculations have compromised the efficiency of DL models in extracting landslides from a large set of RS images. Lightweight convolutional neural networks (CNNs) exhibit promising feature representation abilities with fewer parameters. This study aims to introduce a new lightweight CNN called MS2LandsNet, designed to detect landslides with both high efficiency and accuracy. The MS2LandsNet consists of three down-sampling stages embedded with multi-scale feature fusion (MFF), aiming to decrease parameters while aggregating contextual features. Additionally, we incorporate multi-scale channel attention (MSCA) into MFF to improve performance. According to experimental results on three landslip datasets, MS2LandsNet obtains the highest F1 score of 85.90% and the highest IoU of 75.28%. Notably, MS2LandsNet accomplishes the resuts with the fewest parameters and the fastest inference speed, outperforming seven classical semantic segmentation models and three lightweight CNNs. The proposed lightweight model holds potential for application on a cloud computing platform for larger-scale landslide mapping tasks in future work.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
11  
ISSN:  
1753-8947  
eISSN:  
1753-8955  
入藏号:  
WOS:001137114700001  
Indexed Date:  
2024-01-12  
Institution Address:  
Chinese Acad Sci, Inst Geog Sci & Nat Resources Res, State Key Lab Resources & Environm Informat Syst, Beijing, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Coll Resources & Environm, Beijing, Peoples R China  
Institution Address:  
Northwest A&F Univ, Coll Water Resources & Architectural Engn, Yangling, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Geog Sci & Nat Resources Res, State Key Lab Resources & Environm Informat Syst, Beijing 100101, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Coll Resources & Environm, Beijing 100049, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 72. Multi-manipulation Detection Network Combining Multi-scale Feature and

摘要: With the continuous development of image editing technologies, it is particularly significant to develop image forensics technologies for image content security. Most existing forensics methods concentrated on single image manipulation detection but with weak robustness and no considerations on tampering location. This paper presents a multi-manipulation image forgery detection method based on convolutional neural network. In this network, a convolution flow based on residual block is constructed to extract manipulation features. Then, a multi-scale feature fusion module is designed to achieve operational feature fusion at different scales. Finally, the fused manipulation features are fed into the multi-branch prediction module, predicting the type and location of each utilized manipulation as the multi-manipulation detection results. An image dataset produced by multiple typical image manipulations is built to train and test the proposed network. The experimental results show that the proposed scheme can recognize the type of tampered manipulations and locate the tampered area more accurately with fewer parameters, and has better robustness to common image post-processing operations, compared with the state-of-the-art object detection networks.  
摘要:  
面对不断进步的图像编辑技术,发展相应的图像取证技术显得尤为重要.针对现有图像篡改检测技术中存在的可检测操作类型单一、鲁棒性不强、篡改区域定位不足等问题,提出一种基于卷积神经网络的多操作图像篡改检测方案.在该网络中,通过构造基于残差块的卷积流以提取操作特征.然后,设计一个多尺度特征融合模块,实现不同尺寸的操作特征融合.最后,将融合后的操作特征输入多分支预测模块进行篡改类型预测与定位,得到多操作检测结果.本文制作了多操作图像篡改数据集,对提出的网络模型进行训练和测试.实验结果表明,本文方案与主流的目标检测网络相比,能够更准确地对篡改区域进行定位,参数量更少,且对常见的图像后处理具有更好的鲁棒性.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1674-2974  
入藏号:  
CSCD:7523309  
Indexed Date:  
2023-11-10  
Institution Address:  
天津大学电气自动化与信息工程学院, 天津 300072, 中国  
Institution Address:  
北京国网信通埃森哲信息技术有限公司, 北京 100052, 中国  
Institution Address:  
School of Electrical and Information Engineer,Tianjin University, Tianjin 300072, China  
Institution Address:  
Beijing SGITG-ACCENTURE Information Technology Co. Ltd., Beijing 100052, China  
Institution:  
天津大学电气自动化与信息工程学院  
Institution:  
北京国网信通埃森哲信息技术有限公司  
Institution:  
School of Electrical and Information Engineer,Tianjin University  
Institution:  
Beijing SGITG-ACCENTURE Information Technology Co. Ltd.

### 73. Research on a Lightweight Multi-Scale Feature Fusion and its Fault

摘要: Convolutional neural networks(CNNs) show significant potential for bearing fault diagnosis. However, traditional CNNs face challenges such as poor noise resistance, high computational complexity, reliance on extensive samples, and limited generalizability. As a result, this paper proposes WDSC-Net, a lightweight, multiscale feature fusion method, focusing on limited labeled fault samples. Initially, a wide kernel convolutional is employed, aiming to reduce parameters and computational complexity. Next, features are fed into a 1x1 convolutional layer reduces feature dimensionality. Subsequently, leveraging the benefits of depth-separable convolution (DSC) allows the separation of spatial and channel features, constructing four convolutional layers of varying scales to amplify the nonlinear fault representation. Finally, an improved feature soft-threshold denoising module is introduced for global feature denoising. Validation on CWRU and MCDS datasets shows that the WDSC-Net method exhibits superior generalizability and noise resistance compared to typical deep-learning fault methods.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1507-2711  
入藏号:  
WOS:001392013900017  
Indexed Date:  
2025-01-13  
Institution Address:  
Northeastern Univ, Sch Mech Engn & Automat, Shenyang, Peoples R China  
Institution Address:  
Northeastern Univ, Key Lab Vibrat & Control Aeroprop Syst, Minist Educ, Shenyang 110819, Peoples R China  
Institution Address:  
Northeastern Univ, Liaoning Prov Key Lab Multidisciplinary Design Opt, Shenyang 110819, Peoples R China  
Institution Address:  
Northeast Forestry Univ, Coll Mech & Elect Engn, Harbin, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 74. Enhanced Multiscale Feature Fusion Network for HSI Classification

摘要: Deep learning-based hyperspectral image (HSI) classification methods have recently attracted significant attention. However, features captured by convolutional neural network (CNN) are always partial due to the restrictions of the respective fields and the loss of multiscale information, which lead to features being discontinuous when extracted. In a departure from existing approaches, in this article, we propose a novel Enhanced Multiscale Feature Fusion Network (EMFFN). As a deeper and wider network, EMFFN can extract sufficiently multiscale features from the parallel multipath of three stages for HSI classification purposes. There are two subnetworks for multiscale spectral and spatial information in EMFFN, respectively. First, we propose a spectral Cascaded Dilated Convolutional Network (CDCN) designed to obtain a larger respective field for long-ranged information and extract multiscale features. Subsequently, a Parallel Multipath Network (PMN) is proposed to capture large-scale, middle-scale, and small-scale spatial features in parallel during all three stages. In the next step, hierarchical features are fused successively, and shallower feature maps can achieve better learning performance when guided by deeper semantic information. As PMN deepens in different stages, more multiscale information flows into the network, enabling finer classification results. To incorporate abundant spectral and spatial features, moreover, we combine features collected from two subnetworks into EMFFN using the designed consolidated loss function. As a result, the network facilitates the learning of not only localization-preserved features, but also high-level semantic features. In our experiments, three benchmark HSIs are utilized to evaluate the performance of the proposed method. Our results demonstrate that the proposed EMFFN can outperform state-of-the-art methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
56  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
被引频次合计:  
57  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000722170500042  
Indexed Date:  
2021-12-05  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & R, Wuhan 430079, Peoples R China  
Institution Address:  
Wuhan Univ, Natl Engn Res Ctr Multimedia Software, Inst Artificial Intelligence, Sch Comp Sci, Wuhan 430072, Peoples R China  
Institution Address:  
Wuhan Univ, Hubei Key Lab Multimedia & Network Commun Engn, Wuhan 430072, Peoples R China  
Institution:  
Institution:  
Institution:

### 75. GLIC: Underwater target detection based on global-local information

摘要: With the rapid development of object detection technology, underwater object detection has attracted widespread attention. Most of the existing underwater target detection methods are built based on convolutional neural networks (CNNs), which still have some limitations in the utilization of global information and cannot fully capture the key information in the images. To overcome the challenge of insufficient global-local feature extraction, an underwater target detector (namely GLIC) based on global-local information coupling and multiscale feature fusion is proposed in this paper. Our GLIC consists of three main components: spatial pyramid pooling, global-local information coupling, and multi-scale feature fusion. Firstly, we embed spatial pyramid pooling, which improves the robustness of the model while retaining more spatial information. Secondly, we design the feature pyramid network with global-local information coupling. The global context of the transformer branch and the local features of the CNN branch interact with each other to enhance the feature representation. Finally, we construct a Multi-scale Feature Fusion (MFF) module that utilizes balanced semantic features integrated at the same depth for multi-scale feature fusion. In this way, each resolution in the pyramid receives equal information from others, thus balancing the information flow and making the features more discriminative. As demonstrated in comprehensive experiments, our GLIC, respectively, achieves 88.46%, 87.51%, and 74.94% mAP on the URPC2019, URPC2020, and UDD datasets.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1047-3203  
eISSN:  
1095-9076  
入藏号:  
WOS:001357220200001  
Indexed Date:  
2024-11-24  
Institution Address:  
Dalian Maritime Univ, Coll Marine Elect Engn, Dalian, Peoples R China  
Institution Address:  
Chongqing Univ, Key Lab Optoelect Technol & Syst, Educ Minist, Chongqing, Peoples R China  
Institution:  
Institution:

### 76. Global-Local Transformer Network for HSI and LiDAR Data Joint

摘要: Hyperspectral images (HSIs) contain rich spatial and spectral detail information, while light detection and ranging (LiDAR) data can provide the elevation information. Thus, the fusion of HSI and LiDAR data can help in more accurate image classification, which becomes a hot research topic. However, it is difficult to capture complex local and global spatial-spectral associations; meanwhile, how to build an effective interaction between multimodal data is another important issue. To this end, a novel global-local transformer network (GLT-Net) is proposed for the joint classification of HSI and LiDAR data, in this article. The main idea is to fully exploit the advantage of the convolution operator in characterizing locally correlated features and the promising capability of transformer architecture in learning longrange dependencies. Moreover, multiscale feature fusion and probabilistic decision fusion strategies are also designed in one framework, to further improve the classification performance. Here, the proposed GLT-Net mainly consists of multiscale local spatial feature learning, global spectral feature learning, and global-local feature fusion classification. In specific, multimodal image cubes of different sizes are first extracted and sent into convolutional neural networks (CNNs) to learn local spatial features, which is followed by multimodal information propagation and spatial-attention-guided multiscale feature fusion. Afterward, by considering spectral feature channels from a sequential perspective, vision transformers are introduced to model the global spectral dependencies. Finally, multiple class estimations based on local and global features are integrated via a probabilistic decision fusion strategy. In this way, complementary information of multimodal data and local/global spectral-spatial information can be fully mined and jointly used. Extensive experiments on three popular HSI and LiDAR datasets demonstrate that the proposed method performs superiority over the state-of-the-art methods. The source code of the proposed method will be made publicly available at https://github.com/Ding-Kexin/GLT-Net.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
71  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
71  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000880753700018  
Indexed Date:  
2022-12-04  
Institution Address:  
Hunan Univ, Coll Elect & Informat Engn, Changsha 410082, Peoples R China  
Institution Address:  
Hunan Univ, Key Lab Visual Percept & Artificial Intelligence, Changsha 410082, Peoples R China  
Institution:  
Institution:

### 77. 3D Object detector: A multiscale region proposal network based on

摘要: Recently, 3D object detection by point cloud data processing has been applied to robotics and autonomous driving because of the popularity of the LiDAR sensors. Point cloud data contain the depth and geometric space information of an object as compared with the 2D images, and achieve high precision for classification and location. In the traditional processing of point cloud data, the disorder and sparsity of the points are significant problems. In addition, the traditional detector can only support processing a limited number of point clouds. Thus, it is difficult to detect objects using a large number of point clouds. However, the previous methods need to sample the point cloud data into a coarser type, so they cannot avoid the loss of information and the accuracy is affected, as seen in in PV-RCNN. In this paper, we propose a multiscale feature fusion detector called multiscale region proposal networks (MS-RPNs), which can provide multiscale prediction results for difficult category objects. Meanwhile, our method can improve the detection accuracy for smaller objects with the optimal processing of the multiscale feature extraction module. The efficiency and accuracy of the multiscale region proposal network on the KITTI 3D object detection datasets was evaluated using numerous experiments.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0045-7906  
eISSN:  
1879-0755  
入藏号:  
WOS:000879218900003  
Indexed Date:  
2022-11-18  
Institution Address:  
Yangzhou Univ, Yangzhou, Peoples R China  
Institution Address:  
Qingdao Univ, Qingdao, Peoples R China  
Institution Address:  
Kyushu Inst Technoligy, Kitakyushu, Japan  
Institution Address:  
Univ Toronto, Toronto, ON, Canada  
Institution:  
Institution:  
Institution:  
Kyushu Inst Technoligy  
Institution:

### 78. EMC<SUP>2</SUP>A-Net: An Efficient Multibranch Cross-Channel Attention

摘要: In recent years, convolutional neural networks (CNNs) have demonstrated significant potential for synthetic aperture radar (SAR) target recognition. SAR images possess a strong sense of granularity and contain texture features of varying scales, including speckle noise, dominant scatterers, and target contours, which are not typically considered in traditional CNN models. This article proposes two residual blocks, termed multibranch cross-channel attention (EMC(2)A) blocks, with multiscale receptive fields (RFs) based on a multibranch structure and designs an efficient isotopic architecture deep CNN (DCNN) called EMC(2)A-Net, whose structure is interpretable from a probability and mathematical statistics perspective. EMC(2)A blocks employ parallel dilated convolution with different dilation rates to effectively capture multiscale contextual features without significantly increasing the computational load. To further enhance the efficiency of multiscale feature fusion, this article presented a multiscale feature cross-channel attention module, known as the EMC(2)A module, which adopts a local multiscale feature interaction strategy without dimensionality reduction. This strategy adaptively adjusts the weights of each channel using efficient one-dimensional (1-D)-circular convolution and sigmoid function to guide attention at the global channel-wise level. Comparative results on the moving and stationary target acquisition and recognition (MSTAR) dataset demonstrate that EMC(2)A-Net outperforms the other available models of the same type and possesses a relatively lightweight network structure. The ablation experimental results further demonstrate that the EMC(2)A module significantly enhances the model's performance by utilizing only a few parameters and appropriate cross-channel interactions.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001022708100022  
Indexed Date:  
2023-08-05  
Institution Address:  
Nanjing Inst Technol, Sch Comp Engn, Nanjing 211167, Peoples R China  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Elect & Informat Engn, Nanjing 211106, Peoples R China  
Institution:  
Institution:

### 79. CMFuse: Correlation-based multi-scale feature fusion network for the

摘要: COVID-19 broke out in 2019, seriously affecting people's health and life. Recent studies have indicated that radiological images carry crucial information about COVID-19. Hence, automatic image classification assisted by artificial intelligence (AI) can be employed as a potential diagnostic tool. Nonetheless, in the task of COVID-19 X-ray image recognition, there are local features, including local vascular dilatation, as well as global features, including large ground glass-like shadows, traditional deep neural networks cannot effectively extract features, and the significance of distinct scale features for the task is also divergent, feature element-wise adding or feature concatenating to fuse features from various branches do not consider the internal correlation between features. In view of the above problems, we propose a Correlation-based Multi-scale Feature Fusion Network (CMFuse), combining the advantages of Convolutional Neural Network (CNN) and Transformer. The model captures local spatial contextual features and global semantic information representation of features at different scales in parallel, and the extracted features are adaptively fused at distinct levels through the feature fusion module, down-sampling and other steps to obtain the final classification results. We evaluated CMFuse on the integrated COVID-19 X-ray image dataset, and the results showed that our model attains 97.36% Accuracy, 99.15% Specificity, 97.27% Recall, 97.17% Precision, and 97.22% F1-score, which outperforms other previous related works.  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001090305500003  
Indexed Date:  
2023-11-18  
Institution Address:  
China Jiliang Univ, Coll Informat Engn, Key Lab Electromagnet Wave Informat Technol & Met, Hangzhou, Zhejiang, Peoples R China  
Institution Address:  
Hangzhou Canc Hosptial, Dept Pathol, Hangzhou, Zhejiang, Peoples R China  
Institution Address:  
Stevens Inst Technol, Dept Elect & Comp Engn, Hoboken, NJ USA  
Institution:  
Institution:  
Hangzhou Canc Hosptial  
Institution:

### 80. Low-speed impact localization of wind turbine blades with a single

摘要: Impact, which may occur during manufacturing, serving and maintaining, is a significant threat to in-service composite structures, e.g. wind turbine blades. It calls for developing a method for assessment and localization of impact. In this paper, a single-sensor impact localization method based on deep learning is proposed. Specifically, a multiscale feature fusion convolutional neural network is designed, which, in combination with a convolutional block attention module, adaptively extracts features from single-sensor signals to achieve accurate region-level source localization. Complete ensemble empirical mode decomposition with adaptive noise is employed to reduce noise and extract intrinsic mode functions from acoustic emission signals, enabling more effective feature extraction. The decomposed signals are then converted into grayscale images, forming a dataset for the deep learning model. This approach allows for the extraction of rich feature information. A steel ball drop experiment is conducted to simulate the low-speed impact response of the wind turbine blade spar. The experimental results show significant advantages in localization accuracy. This study offers a promising solution for acoustic emission source region localization in complex composite structures.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0041-624X  
eISSN:  
1874-9968  
入藏号:  
WOS:001440074000001  
PubMed ID:  
39955861  
Indexed Date:  
2025-03-15  
Institution Address:  
Xi An Jiao Tong Univ, Sch Mech Engn, Xian 710049, Shaanxi, Peoples R China  
Institution Address:  
713th Res Inst China Shipbuilding Ind Corp, Zhengzhou 450015, Henan, Peoples R China  
Institution:  
Institution:  
713th Res Inst China Shipbuilding Ind Corp

### 81. MLFFNet: Multilevel Feature Fusion Network for Object Detection in Sonar

摘要: Sonar image object detection is essential in underwater rescue and resource exploration. Although many convolution neural network (CNN)-based object detection algorithms have achieved great success in natural images. However, for underwater sonar images, problems, such as seabed reverberation noise interference, low proportion of foreground object region pixels, and poor imaging resolution, present considerable challenges to achieving accurate underwater object detection. To address these problems, we propose a novel sonar image object detector called the multilevel feature fusion network (MLFFNet). The detector consists of multiscale convolution module (MS-Conv), multilevel feature extraction module (ML-FEM), multilevel feature fusion module (ML-FFM), neighborhood channel attention mechanism (N-CAM), multiscale feature pyramid module (MS-FPN), and feature association module (FA). First, we use the MS-Conv to extract different scale feature information in the object region. Second, the ML-FEM and ML-FFM are used to obtain the local detail and global context features. Third, the N-CAM and MS-FPN are used to obtain the foreground objects' semantic feature and position feature, and suppress the background region noise interference. Finally, we use the FA module to enhance the category and feature correlation of different objects. Extensive experiments are conducted on the real scene sonar image dataset. The experimental results demonstrate that MLFFNet performs better than other state-of-the-art object detection methods. Code and dataset are publicly at https://github.com/darkseid-arch/SonarMLFFNet.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
19  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
19  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000878208400006  
Indexed Date:  
2022-11-15  
Institution Address:  
Xijing Univ, Sch Informat Engn, Xian 710123, Peoples R China  
Institution Address:  
Air Force Engn Univ, Sch Informat & Nav, Xian 710082, Peoples R China  
Institution Address:  
Tianjin Univ Sci & Technol, Coll Artificial Intelligence, Tianjin 300457, Peoples R China  
Institution:  
Institution:  
Institution:

### 82. An Improved 3D-2D Convolutional Neural Network Based on Feature

摘要: As a new technology in the field of remote sensing, hyperspectral remote sensing has been widely used in land classification, mineral exploration, environmental monitoring, and other areas. In recent years, deep learning has achieved outstanding results in hyperspectral image classification tasks. However, problems such as low classification accuracy for small sample classes in unbalanced datasets and lack of robustness of the models usually lead to unstable classification performance of hyperspectral images. Therefore, from the perspective of feature optimization, we propose an improved hybrid convolutional neural network for hyperspectral image feature extraction and classification. Different from the current simple multi-scale feature extraction, we first optimize the features of each scale, and then perform multi-scale feature fusion. To this end, we use 3D dilated convolution to design a multi-level feature extraction block (MFB), which can be used to extract features with different correlation strengths at a fixed scale. Then, we construct a spatial multi-scale interactive attention (SMIA) module in the spatial feature enhancement phase, which can refine the multi-scale features through the attention weights of multi-scale feature interaction, and further improve the quality of spatial features. Finally, experiments were performed on different datasets, including balanced and unbalanced samples. The results show that the proposed model is more accurate and the extracted features are more robust.  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
2169-3536  
入藏号:  
WOS:000966308900001  
Indexed Date:  
2023-04-28  
Institution Address:  
Henan Polytech Univ, Sch Surveying & Land Informat Engn, Jiaozuo 454003, Peoples R China  
Institution Address:  
Collaborat Innovat Ctr Geoinformat Technol Smart C, Zhengzhou 450000, Henan, Peoples R China  
Institution Address:  
Minist Nat Resources, Key Lab Spatiotemporal Percept & Intelligent Proc, Zhengzhou 450000, Peoples R China  
Institution:  
Institution:  
Collaborat Innovat Ctr Geoinformat Technol Smart C  
Institution:

### 83. Fabric Defect Classification Algorithm Based on Multi-Scale Feature

摘要: A fabric surface defect classification technique based on spatial attention multiscale feature fusion is designed to address the problem of low-classification accuracy caused by complex texture and varied defect kinds of the fabric surface. The multiscale pyramid pooling module is used to maintain the information integrity of the feature map, and the rich semantic information extracted from the high-level feature map is used as a priori information to guide the low-level features, realizing the fusion of high-level and low-level features; the improved spatial attention module is integrated into a convolutional neural network to enhance the differential expression of features. The improved class activation mapping method is used to obtain the defect classification information and location information. The fabric surface defect image is recognized and detected using data augmentation and transfer learning methods. The experimental results show that the proposed algorithm can effectively increase the accuracy of fabric defect classification and obtain defect location information without manual location labeling.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1006-4125  
入藏号:  
WOS:000823175200011  
Indexed Date:  
2022-07-21  
Institution Address:  
Zhejiang Sci Tech Univ, Sch Mech & Automat Control, Hangzhou 310018, Zhejiang, Peoples R China  
Institution:

### 84. A developed convolutional neural network model for accurately and stably

摘要: Accurate and stable prediction of the effective thermal conductivity (ETC) of porous ceramic materials is of great significance for their application in areas such as optimizing the design of thermal barrier coatings and improving energy conversion efficiency. Porous ceramic materials exhibit complex porous structures with gradient porosity distributions that pose a great challenge for effective medium theory (EMT) and conventional convolutional neural networks (CNN) in attempting to accurately and stably predict the ETC of porous media. In this study, a CNN model that integrates self-attention and a multiscale feature-fusion mechanism is proposed to predict the ETC of porous media with greater accuracy and stability. The integration of the self-attention and multiscale feature-fusion mechanisms enhances the CNN's ability to learn long-range dependencies and preserve detailed information. The optimization of the CNN's accuracy and stability is visually illustrated using gradient-weighted class activation mapping (Grad-CAM) for ETC. Additionally, by employing the proposed gradient quartet structure generation set (QSGS), a gradient porous ceramic media dataset comprising 10 000 images was built to train the CNN model. Finally, the prediction results and relative error distribution of the ETC were compared across different models. Our model demonstrated improvements in statistical metrics, including a 33.7 % decrease in mean error, 25.2 % decrease in median error, and 59.6 % decrease in maximum error. The decreases in these metrics and Grad-CAM for the ETC strongly demonstrates that the model proposed in this work greatly improved the accuracy and stability of predicting the ETC of ceramic materials with gradient porosity distributions.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0017-9310  
eISSN:  
1879-2189  
入藏号:  
WOS:001209638900001  
Indexed Date:  
2024-05-03  
Institution Address:  
Xi An Jiao Tong Univ, Sch Chem Engn & Technol, Xian 710049, Peoples R China  
Institution Address:  
Beijing Inst Astronaut Syst Engn, Beijing 100076, Peoples R China  
Institution Address:  
Xi An Jiao Tong Univ, Sch Elect & Informat Engn, Xian 710049, Peoples R China  
Institution:  
Institution:  
Beijing Inst Astronaut Syst Engn  
Institution:

### 85. E-FPN: Evidential Feature Pyramid Network for Ship Classification

摘要: Ship classification, as an important problem in the field of computer vision, has been the focus of research for various algorithms over the past few decades. In particular, convolutional neural networks (CNNs) have become one of the most popular models for ship classification tasks, especially using deep learning methods. Currently, several classical methods have used single-scale features to tackle ship classification, without paying much attention to the impact of multiscale features. Therefore, this paper proposes a multiscale feature fusion ship classification method based on evidence theory. In this method, multiple scales of features were utilized to fuse the feature maps of three different sizes (40 x 40 x 256, 20 x 20 x 512, and 10 x 10 x 1024), which were used to perform ship classification tasks separately. Finally, the multiscales-based classification results were treated as pieces of evidence and fused at the decision level using evidence theory to obtain the final classification result. Experimental results demonstrate that, compared to classical classification networks, this method can effectively improve classification accuracy.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001045650800001  
Indexed Date:  
2023-08-23  
Institution Address:  
Shanghai Maritime Univ, Coll Informat Engn, Shanghai 201306, Peoples R China  
Institution Address:  
Shanghai Maritime Univ, Coll Logist Engn, Shanghai 201306, Peoples R China  
Institution Address:  
Jiangsu Automat Res Inst, Lianyungang 222061, Peoples R China  
Institution:  
Institution:  
Institution:  
Jiangsu Automat Res Inst

### 86. Ground Object Feature Extraction fromRemote Sensing Images Based on

摘要: The information extraction of ground objects from remote sensing images are of great significance for missile forces to complete the task of long-range precision strikes. At present, semantic segmentation network represented by convolutional neural networks (CNN)can realize the semantic segmentation of daily images beter, it is still difficult to capture global information and long-range dependencies in remote sensing images. To address these problems, an approach based on multi-scale feature fusion and dual attention network is presented for ground object feature extraction of remote sensing images. The method is based on the architecture of the deep encoder-decoder network, a multi-scale feature fusion module and dual-attention network are introduced. The multi-scale feature fusion module can extract the multi-scale features of images and can capture global information of images. The dual-attention module can capture the long-range context relations of images and the interdependence between channels. The experiments on the GID dataset show that the proposed method can achieve better effects of ground object feature extraction compared with the other existing methods, with the overall segmentation accuracy of 84.1% and the mIoU of 62.3%.  
摘要:  
遥感影像的地物信息提取,对导弹部队完成远程精确打击任务具有重要意义。目前,卷积神经网络为代表的语义分割网络能较好实现日常图像的语义分割,但难以捕捉遥感图像的全局信息和长距离依赖关系。提出了一种于多尺度特征融合和双注意力网络的遥感影像地物提取方法。该方法在编解码网络的基础上,引入多尺度特征融合模块和双注意力网络。多尺度融合模块能够提取图像的多尺度特征,捕获图像的全局信息。注意力模块能够捕获图像远程上下文关系和通道间的相互依赖。在GID数据集上的实验结果表明,与现有方法相比,该方法能够获得更好的地物提取效果,其分割的总体精度为84.1%,mIoU为62.3%。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-0640  
入藏号:  
CSCD:7604527  
Indexed Date:  
2024-02-02  
Institution Address:  
火箭军士官学校, 青州, 山东 262500, 中国  
Institution Address:  
PLA Rocket Force NCO College, Qingzhou, 262500  
Institution:  
火箭军士官学校  
Institution:  
PLA Rocket Force NCO College

### 87. MSAFFNet: A Multiscale Label-Supervised Attention Feature Fusion Network

摘要: The detection of small infrared targets with low signal-to-noise ratios (SNRs) and contrasts in noisy and cluttered backgrounds is challenging and therefore a domain of active research. Traditional methods result in a large number of false alarms and missed detections. In the case of convolutional neural network (CNN)-based methods, it may not be possible to identify deep small targets or the details of the target's edge contours may not be appropriately considered. Therefore, this article proposes MSAFFNet to perform infrared small target detection (IRSTD) based on an encoder-decoder framework. In the encoder stage, small target features are extracted using a resnet-20 backbone network, and the global contextual features of small targets are extracted using an atrous spatial pyramid pooling module (ASPPM). In the decoding stage, a dual-attention module (DAM) is used to selectively enhance the spatial details of the target at the shallow level and representative features of the semantic information at the deep level. Multiscale feature maps are then concatenated to achieve superior feature fusion. Additionally, multiscale labels are constructed to focus on the details of the target contour and internal features based on edge information and an internal feature aggregation module (EIFAM). Experiments conducted on the nanjing university of aeronautics and astronautics-single-frame infrared small target (NUAA-SIRST), national university of defense technology- SIRST (NUDT-SIRST), and xidian university-SIRST (XDU-SIRST) datasets revealed that the proposed approach outperforms the representative methods and achieves an improved detection performance.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
28  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
30  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001005737500019  
Indexed Date:  
2023-07-18  
Institution Address:  
Natl Univ Def Technol, Coll Intelligence Sci & Technol, Changsha 410073, Peoples R China  
Institution:

### 88. Assisting RGB and depth salient object detection with nonconvolutional

摘要: RGB-D salient object detection is a challenging task in computer vision, and deep architectures have been widely adopted in the previous studies. However, current convolutional neural network (CNN)-based models struggle with capturing global long-distance features efficiently, whereas transformer-based methods are computationally intensive. To address these limitations, we propose a nonconvolutional feature encoder. This encoder captures long-distance dependencies while reducing computation costs, making it a potential alternative to CNNs and transformers. Additionally, we introduce a spatial info enhancing mechanism to overcome weakened local information while capturing long-range dependencies. This mechanism balances local and global information at different expansion rates by exploring multiscale feature fusion in the feature maps. Furthermore, we introduce a spatial info sensing module to enhance the compatibility of multimodal features in long-range dependencies and extract informative cues from depth features. Through comprehensive experiments on four widely used datasets, we demonstrate that our proposed involution encoder significantly outperforms previous state-of-the-art RGB-D salient object detection methods based on CNNs in four key metrics. Compared to transformer-based methods, our approach balances speed and efficiency favorably. (c) 2024 SPIE and IS&T  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:001286842600011  
Indexed Date:  
2024-08-16  
Institution Address:  
China Univ Petr East China, Coll Comp Sci & Technol, Qingdao, Peoples R China  
Institution:

### 89. Tiny-Crack-Net: A multiscale feature fusion network with attention

摘要: Convolutional neural networks (CNNs) have gained growing interest in recent years for their advantages in detecting cracks on concrete bridge components. Class imbalance is a fundamental problem in crack segmentation, resulting in unsatisfactory segmentation for tiny cracks. Besides, limited by the local receptive field, CNNs often cannot integrate local features with global dependencies, thus significantly affecting the detection accuracy of tiny cracks across the entire image. To solve those problems in segmenting tiny cracks, a multiscale feature fusion network with attention mechanisms named "Tiny-Crack-Net" (TCN) is proposed. The modified residual network was used to capture the local features of tiny cracks. The dual attention module was then incorporated into the architecture to better separate the tiny cracks from the background. Also, a multiscale fusion operation was implemented to preserve the edge details of tiny cracks. Finally, a joint learning loss of the cross-entropy and similarity was proposed to alleviate the poor convergence induced by the severe class imbalance of the pixels representing tiny cracks. The capability of the network in segmenting tiny cracks was remarkably enhanced by the aforementioned arrangements, and the "Tiny-Crack-Net" achieved a Dice similarity coefficient of 87.96% on an open-source data set, which was at least 5.84% higher than those of the six cutting-edge networks. The effectiveness and robustness of the "Tiny-Crack-Net" were validated with field test results, which showed that the intersection over union (IOU) for cracks with a width of 0.05 mm or wider reaches 91.44%.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
85  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
88  
ISSN:  
1093-9687  
eISSN:  
1467-8667  
入藏号:  
WOS:000815587900001  
Indexed Date:  
2022-07-01  
Institution Address:  
Hunan Univ, Coll Civil Engn, Changsha 410082, Peoples R China  
Institution Address:  
Hunan Univ, Key Lab Damage Diag Engn Struct Hunan Prov, Changsha, Peoples R China  
Institution:  
Institution:

### 90. Human centric attention with deep multiscale feature fusion framework

摘要: Recent advancements in the Internet of Medical Things (IoMT) have revolutionized the healthcare sector, making it an active research area in the academic and industrial sectors. Following these advances, an automatic Human Activity Recognition (HAR) is now integrated into the IoMT, facilitating remote patient monitoring systems for smart healthcare. However, implementing HAR via computer vision is intricate due to complex spatiotemporal patterns, single stream fusion, and clutter backgrounds. Mainstream approaches practice pre-trained CNN model, which extract non-salient features due to their generalized weight optimization and limited discriminative feature fusion. In addition, their sequential models have inadequate performance in complex scenarios due to the vanishing gradients encountered during backpropagation across multiple layers. In response to these challenges, we propose a multiscale feature fusion framework for both indoor and outdoor environments to enhance HAR in healthcare monitoring systems, which is mainly composed of two stages: First, the proposed Human Centric Attentional Fusion (HCAF) network is fused with the intermediate convolutional feature of lightweight MobileNetV3 backbone to enriches spatial learning capabilities for accurate HAR. Next, a Deep Multiscale Features Fusion (DMFF) network is proposed that enhanced the long-range temporal dependencies by redesigning the traditional bidirectional LSTM network into a residual fashion followed by Sequential Multihead Attention (SMA) to eliminate non-relevant information and optimized spatiotemporal feature vectors. The performance of the proposed fusion model is evaluated on benchmark healthcare and general activity datasets. In the healthcare, we used Multiple Camera Fall and UR Fall Detection datasets that achieved 99.941% and 100% accuracy. Despite this, our fusion strategy is rigorously evaluated over three challenging general HAR datasets, including HMDB51, UCF101, and UCF50, demonstrating 74.942%, 97.337%, and 96.156% superior performance compared to Stateof-The-Art (SOTA) methods. The run time analysis shows that the proposed method is 2x times faster than the existing methods.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
ISSN:  
1566-2535  
eISSN:  
1872-6305  
入藏号:  
WOS:001184568200001  
Indexed Date:  
2024-03-29  
Institution Address:  
Sejong Univ, Seoul 143747, South Korea  
Institution Address:  
Univ Technol Compiegne, Ctr Rech Royallieu, Lab Biomecan & Bioingn, Compiegne, France  
Institution:  
Institution:

### 91. AFSRNet: learning local descriptors with adaptive multi-scale feature

摘要: Multi-scale feature fusion has been widely used in handcrafted descriptors, but has not been fully explored in deep learning-based descriptor extraction. Simple concatenation of descriptors of different scales has not been successful in significantly improving performance for computer vision tasks. In this paper, we propose a novel convolutional neural network, based on center-surround adaptive multi-scale feature fusion. Our approach enables the network to focus on different center-surround scales, resulting in improved performance. We also introduce a novel regularization technique that uses second-order similarity to constrain the learning of local descriptors, based on the symmetric property of the similarity matrix. The proposed method outperforms single-scale or simple-concatenation descriptors on two datasets and achieves state-of-the-art results on the Brown dataset. Furthermore, our method demonstrates excellent generalization ability on the HPatches dataset. Our code is released on GitHub: https://github.com/Leung-GD/AFSRNet/tree/main.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:001205422800001  
Indexed Date:  
2024-04-28  
Institution Address:  
Guangdong Univ Technol, Guangzhou 510006, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Hong Kong 999077, Peoples R China  
Institution:  
Institution:

### 92. Ship Detection in SAR Images Based on Multiscale Feature Fusion and

摘要: Deep-learning technology has enabled remarkable results for ship detection in SAR images. However, in view of the complex and changeable backgrounds of SAR ship images, how to accurately and efficiently extract target features and improve detection accuracy and speed is still a huge challenge. To solve this problem, a ship detection algorithm based on multiscale feature fusion and channel relation calibration of features is proposed in this paper. First, based on Faster R-CNN, a channel attention mechanism is introduced to calibrate the channel relationship between features in the feature extraction network, so as to improve the network's expression ability for extraction of ship features in different scenes. Second, unlike the original method of generating candidate regions based on single-scale features, this paper introduces an improved feature pyramid structure based on a neural architecture search algorithm, which helps improve the performance of the network. The multiscale features are effectively fused to settle the problem of missing detections of small targets and adjacent inshore targets. Experimental results on the SSDD dataset show that, compared with the original Faster R-CNN, the proposed algorithm improves detection accuracy from 85.4% to 89.4% and the detection rate from 2.8 FPS to 10.7 FPS. Thus, this method effectively achieves high-speed and high-accuracy SAR ship detection, which has practical benefits.  
摘要:  
目前深度学习技术在SAR图像的船舶检测中已取得显著的成果,但针对SAR船舶图像中复杂多变的背景环境,如何准确高效地提取目标特征,提升检测精度与检测速度仍存在着巨大的挑战。针对上述问题,该文提出了一种多尺度特征融合与特征通道关系校准的SAR图像船舶检测算法。在Faster R-CNN的基础上,首先通过引入通道注意力机制对特征提取网络进行特征间通道关系校准,提高网络对复杂场景下船舶目标特征提取的表达能力;其次,不同于原始的基于单一尺度特征生成候选区域的方法,该文基于神经架构搜索算法引入改进的特征金字塔结构,高效地将多尺度特征进行充分融合,改善了船舶目标中对小目标、近岸密集目标的漏检问题。最后,在SSDD数据集上进行对比验证。实验结果表明,相较原始的Faster R-CNN,检测精度从85.4%提高到89.4%,检测速率也从2.8 FPS提高到10.7 FPS。该方法能够有效实现高速与高精度的SAR图像船舶检测,具有一定的现实意义。  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
8  
ISSN:  
2095-283X  
入藏号:  
CSCD:7038149  
Indexed Date:  
2021-11-19  
Institution Address:  
中国科学院大学;;中国科学院空天信息创新研究院, ;;, ;;, 北京;;北京 100049;;100190, 中国  
Institution Address:  
中国科学院空天信息创新研究院, 北京 100190, 中国  
Institution Address:  
University of Chinese Academy of Sciences;;Aerospace Information Research Institute,Chinese Academy of Sciences, ;;, ;;, Beijing;;Beijing 100049;;100190  
Institution Address:  
Aerospace Information Research Institute,Chinese Academy of Sciences, Beijing 100190, China  
Institution:  
中国科学院大学;;中国科学院空天信息创新研究院  
Institution:  
中国科学院空天信息创新研究院  
Institution:  
University of Chinese Academy of Sciences;;Aerospace Information Research Institute,Chinese Academy of Sciences  
Institution:  
Aerospace Information Research Institute,Chinese Academy of Sciences

### 93. A Real-Time Bridge Crack Detection Method Based on an Improved

摘要: Bridge crack detection is essential to ensure bridge safety. The introduction of deep learning technology has made it possible to detect bridge cracks automatically and accurately. In this study, the Inception-Resnet-v2 algorithm was systematically improved and applied to the real-time detection of bridge cracks. We propose an end-to-end bridge crack detection model based on a convolutional neural network. This model combines the advantages of Inception convolution and residual networks, broadening the network width and alleviating the training problem of the deep network. The calculation speed is improved while still ensuring accuracy. Multi-scale feature fusion enables the network to extract contextual information of different scales, which improves the accuracy of crack recognition. The GKA (K-means clustering method based on a genetic algorithm) realizes the accurate segmentation of the target area, greatly enhances the clustering effect, and effectively improves the detection speed. In this model, large fracture datasets are used for training and testing without pre-training. The experimental results show that the performance of this method was improved in all aspects: accuracy, 99.24%; recall, 99.03%; F-measure, 98.79%; and FPS(Frames Per Second), 196.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
37  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
38  
ISSN:  
2169-3536  
入藏号:  
WOS:000673946600001  
Indexed Date:  
2021-07-27  
Institution Address:  
Army Engn Univ PLA, Coll Field Engn, Dept Mech Engn, Nanjing, Peoples R China  
Institution:

### 94. AFMSFFNet: An Anchor-Free-Based Feature Fusion Model for Ship Detection

摘要: This paper aims to improve a small-scale object detection model to achieve detection accuracy matching or even surpassing that of complex models. Efforts are made in the module design phase to minimize parameter count as much as possible, thereby providing the potential for rapid detection of maritime targets. Here, this paper introduces an innovative Anchor-Free-based Multi-Scale Feature Fusion Network (AFMSFFNet), which improves the problems of missed detection and false positives, particularly in inshore or small target scenarios. Leveraging the YOLOX tiny as the foundational architecture, our proposed AFMSFFNet incorporates a novel Adaptive Bidirectional Fusion Pyramid Network (AB-FPN) for efficient multi-scale feature fusion, enhancing the saliency representation of targets and reducing interference from complex backgrounds. Simultaneously, the designed Multi-Scale Global Attention Detection Head (MGAHead) utilizes a larger receptive field to learn object features, generating high-quality reconstructed features for enhanced semantic information integration. Extensive experiments conducted on publicly available Synthetic Aperture Radar (SAR) image ship datasets demonstrate that AFMSFFNet outperforms the traditional baseline models in detection performance. The results indicate an improvement of 2.32% in detection accuracy compared to the YOLOX tiny model. Additionally, AFMSFFNet achieves a Frames Per Second (FPS) of 78.26 in SSDD, showcasing superior efficiency compared to the well-established performance networks, such as faster R-CNN and CenterNet, with efficiency improvement ranging from 4.7 to 6.7 times. This research provides a valuable solution for efficient ship detection in complex backgrounds, demonstrating the efficacy of AFMSFFNet through quantitative improvements in accuracy and efficiency compared to existing models.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001323792300001  
Indexed Date:  
2024-10-07  
Institution Address:  
Xidian Univ, Sch Phys, Xian 710071, Peoples R China  
Institution:

### 95. RCAG-Net: Residual Channelwise Attention Gate Network for Hot Spot

摘要: The small hot spot defect detection for photovoltaic (PV) farms is a challenging problem due to the feature vanishing as the network deepens. To solve this challenging problem, a novel residual channelwise attention gate network (RCAG-Net) is proposed by employing a novel RCAG module to achieve multiscale feature fusion, complex background suppression, and defect feature highlighting. In RCAG-Net, the novel RCAG module first realizes feature fusion by adding the features of different scale layers. Next, global average pooling (GAP) and multilayer perceptron (MLP) are used to dimension reduction and refinement of the fused features, then yielding an attention map for channelwise feature reweighting by gate mechanism, which employs selective transmission of the convolution neural network (CNN)-extracted features to achieve informative feature filtering. Moreover, residual connection from the fused features to the final output facilitates the insertion of the new RCAG into some classical pretrained models, without breaking its initial behavior. Finally, the proposed approach is validated through a real defect detection system, and the experimental result clearly verifies its effectiveness for small hot spot detection of PV farms.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
47  
在中国科学引文数据库中的被引频次:  
6  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
53  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000702380100018  
Indexed Date:  
2021-10-13  
Institution Address:  
Hebei Univ Technol, Sch Artificial Intelligence & Data Sci, Tianjin 300130, Peoples R China  
Institution Address:  
Beihang Univ, Sch Comp, Beijing 100191, Peoples R China  
Institution:  
Institution:

### 96. Improved YOLOv8-Seg Based on Multiscale Feature Fusion and Deformable

摘要: Laser-targeted weeding methods further enhance the sustainable development of green agriculture, with one key technology being the improvement of weed localization accuracy. Here, we propose an improved YOLOv8 instance segmentation based on bidirectional feature fusion and deformable convolution (BFFDC-YOLOv8-seg) to address the challenges of insufficient weed localization accuracy in complex environments with resource-limited laser weeding devices. Initially, by training on extensive datasets of plant images, the most appropriate model scale and training weights are determined, facilitating the development of a lightweight network. Subsequently, the introduction of the Bidirectional Feature Pyramid Network (BiFPN) during feature fusion effectively prevents the omission of weeds. Lastly, the use of Dynamic Snake Convolution (DSConv) to replace some convolutional kernels enhances flexibility, benefiting the segmentation of weeds with elongated stems and irregular edges. Experimental results indicate that the BFFDC-YOLOv8-seg model achieves a 4.9% increase in precision, an 8.1% increase in recall rate, and a 2.8% increase in mAP50 value to 98.8% on a vegetable weed dataset compared to the original model. It also shows improved mAP50 over other typical segmentation models such as Mask R-CNN, YOLOv5-seg, and YOLOv7-seg by 10.8%, 13.4%, and 1.8%, respectively. Furthermore, the model achieves a detection speed of 24.8 FPS on the Jetson Orin nano standalone device, with a model size of 6.8 MB that balances between size and accuracy. The model meets the requirements for real-time precise weed segmentation, and is suitable for complex vegetable field environments and resource-limited laser weeding devices.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
eISSN:  
2076-3417  
入藏号:  
WOS:001254402500001  
Indexed Date:  
2024-07-02  
Institution Address:  
Guizhou Univ, Coll Big Data & Informat Engn, Guiyang 550025, Peoples R China  
Institution:

### 97. Detection and Tracking of Low-Altitude Unmanned Aerial Vehicles Based on

摘要: With the popularization of nonmilitary unmanned aerial vehicles (UAVs), UAV-detection technology has become a hotspot in security research. This study proposes a low-altitude UAV-detection and -tracking method based on the optimized YOLOv4. This method combines detection technology based on convolutional neural networks with a tracking algorithm for the first time to achieve dynamic detection of low-altitude UAVs. First, the original YOLO network structure is optimized based on multiscale feature fusion. Thereafter, in combination with the DeepSORT multitarget tracking algorithm, the detection and tracking model is constructed. Training and comparative experiments are performed on the self-built LARotorcraft dataset. The experimental results show that the proposed model can effectively reduce the miss detection rate for small targets. Good real-time performance is obtained with an average detection accuracy of up to 77.2%, and stable tracking of visual targets is realized.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1006-4125  
入藏号:  
WOS:000823188100040  
Indexed Date:  
2022-08-10  
Institution Address:  
Nankai Univ, Coll Elect Informat & Opt Engn, Engn Res Ctr Thin Film Optoelect Technol, Minist Educ, Tianjin 300350, Peoples R China  
Institution Address:  
Tianjin Key Lab Optoelect Sensor & Sensing Networ, Tianjin 300350, Peoples R China  
Institution:  
Institution:  
Tianjin Key Lab Optoelect Sensor & Sensing Networ

### 98. Remote Sensing Image Change Detection Based on Lightweight Transformer

摘要: As deep learning demonstrates excellent performance in remote sensing image change detection (CD), early methods that mostly used convolutional neural networks (CNNs) have limitations in the accuracy due to their insufficient global feature representation, an inherent shortcoming of CNNs. The lack of global feature can lead to notable issues, such as the inability to detect small targets and loss of edge information. In recent years, vision transformers have been employed in CD owing to their powerful global feature representation capabilities. However, pure transformer methods lack effective local feature extraction, which also restricts the performance of CD, while the original transformer models require a large amount of computing resources. To address these issues and improve CD performance, we propose a lightweight transformer-based multiscale feature fusion network. By integrating CNN structures both before and after the multihead self-attention in each layer of the main backbone, we enhance the encoder's local feature extraction ability and reduce the computational complexity through convolution and linear operations. For the siamese encoding outputs at different scales, we design two distinct fusion modules based on depthwise convolution for bitemporal information fusion in deep layers and shallow layers, respectively. Our model employs a multilayer cascaded structure with a deep supervision strategy applied to multiple outputs. Experiments on four public CD datasets demonstrate that our network achieves better performance while maintaining relatively smaller computational complexity compared to other state-of-the-art methods for CD.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001432389100012  
Indexed Date:  
2025-03-06  
Institution Address:  
Xinjiang Univ, Sch Comp Sci & Technol, Urumqi 830046, Peoples R China  
Institution:

### 99. MFF-YOLO: An Accurate Model for Detecting Tunnel Defects Based on

摘要: Tunnel linings require routine inspection as they have a big impact on a tunnel's safety and longevity. In this study, the convolutional neural network was utilized to develop the MFF-YOLO model. To improve feature learning efficiency, a multi-scale feature fusion network was constructed within the neck network. Additionally, a reweighted screening method was devised at the prediction stage to address the problem of duplicate detection frames. Moreover, the loss function was adjusted to maximize the effectiveness of model training and improve its overall performance. The results show that the model has a recall and accuracy that are 7.1% and 6.0% greater than those of the YOLOv5 model, reaching 89.5% and 89.4%, respectively, as well as the ability to reliably identify targets that the previous model error detection and miss detection. The MFF-YOLO model improves tunnel lining detection performance generally.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
1424-8220  
入藏号:  
WOS:001036378600001  
PubMed ID:  
37514784  
Indexed Date:  
2023-08-10  
Institution Address:  
North China Univ Water Resources & Elect Power, Sch Elect Engn, Zhengzhou 450045, Peoples R China  
Institution:

### 100. MFEN: Lightweight multi-scale feature extraction super-resolution

摘要: Deep convolutional neural networks (CNN) have achieved remarkable performance in super-resolution (SR) recently. However, deep CNN-based methods are difficult to be utilized in embedded portable device due to their heavy computation and memory consumption. To solve the above problem, we propose an effective lightweight multi-scale feature extraction super-resolution network (MFEN) by constructing multi-scale feature extraction blocks (MFEB), which progressively obtains multi-scale and hierarchical information. In addition, we also propose an efficient progressive feature fusion (PFF) strategy to aggregate multi-scale informative features. Qualitative and quantitative evaluation results on the benchmark datasets show that our designed method can obtain better performance than most state-of-the-art methods. Moreover, the computation complexity and running time of the MFEN are significantly reduced to provide convenience in real-time image processing technology for embedded devices.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0141-9331  
eISSN:  
1872-9436  
入藏号:  
WOS:000870755000004  
Indexed Date:  
2023-01-15  
Institution Address:  
Sichuan Univ, Coll Elect & Informat Engn, Chengdu 610064, Sichuan, Peoples R China  
Institution Address:  
Sichuan Univ, Sch Aeronaut & Astronaut, Chengdu 610065, Sichuan, Peoples R China  
Institution Address:  
Samsung Elect, 129 Samseong ro, Suwon 16677, Gyeonggi Do, South Korea  
Institution:  
Institution:  
Institution:

### 101. Emotion Recognition of Online Education Learners by Convolutional Neural

摘要: At present, the facial expression recognition model in video communication has problems such as weak network generalization ability and complex model structure, which leads to a large amount of computation. Firstly, the Inception architecture is adopted as a design philosophy. The Visual Geometry Group Network (VGGNet) model is improved. Multiscale kernel convolutional layers are constructed to obtain more expressive features. Secondly, the attention mechanism is integrated into a multiscale feature fusion network to form a multiattention mechanism Convolutional Neural Network (CNN) model. Novel spatial and multichannel attention models are designed. The effects of redundant information and noise are reduced. Finally, experiments are carried out on the Fer2013 dataset and the Extended Cohn-Kanade Dataset (CK+) to verify the detection accuracy of the model. The results show that the Delivered Duty Unpaid (DDU) loss can be used for facial expression recognition in complex environments. After the attention module is added, the overall recognition accuracy of the network on Fer2013 and CK+ has been improved to varying degrees. The addition of the channel attention module has a more obvious effect on the recognition accuracy compared with the spatial attention module. The addition of the attention module enables the network to increase the attention to error-prone samples. The improved network model can better extract the key features of facial expressions, enhance the feature discrimination ability, and improve the recognition accuracy of error-prone expressions. The accuracy rate of facial expression recognition with larger movements is over 98%. Facial expressions are an important way of communication between people, and online video has greatly limited this communication method. The proposed CNN model based on multiscale feature fusion will effectively solve these network limitations and have an important and positive impact on future network information exchange.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1687-5265  
eISSN:  
1687-5273  
入藏号:  
WOS:000815046200007  
PubMed ID:  
35720890  
Indexed Date:  
2022-07-06  
Institution Address:  
Fuzhou Univ Int Studies & Trade, Sch Big Data, Fuzhou 350202, Fujian, Peoples R China  
Institution:  
Fuzhou Univ Int Studies & Trade

### 102. Robust Ship Detection in Infrared Images through Multiscale Feature

摘要: The sophistication of ship detection technology in remote sensing images is insufficient, the detection results differ substantially from the practical requirements, mainly reflected in the inadequate support for the differentiated application of multi-scene, multi-resolution and multi-type target ships. To overcome these challenges, a ship detection method based on multiscale feature extraction and lightweight CNN is proposed. Firstly, the candidate-region extraction method, based on a multiscale model, can cover the potential targets under different backgrounds accurately. Secondly, the multiple feature fusion method is employed to achieve ship classification, in which, Fourier global spectrum features are applied to discriminate between targets and simple interference, and the targets in complex interference scenarios are further distinguished by using lightweight CNN. Thirdly, the cascade classifier training algorithm and an improved non-maximum suppression method are used to minimise the classification error rate and maximise generalisation, which can achieve final-target confirmation. Experimental results validate our method, showing that it significantly outperforms the available alternatives, reducing the model size by up to 2.17 times while improving detection performance be improved by up to 5.5% in multi-interference scenarios. Furthermore, the robustness ability was verified by three indicators, among which the F-measure score and true-false-positive rate can increase by up to 5.8% and 4.7% respectively, while the mean error rate can decrease by up to 38.2%.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
eISSN:  
1424-8220  
入藏号:  
WOS:000754571200001  
PubMed ID:  
35161971  
Indexed Date:  
2022-03-04  
Institution Address:  
Beihang Univ, Sch Comp Sci & Engn, Beijing Key Lab Digital Media, Beijing 100191, Peoples R China  
Institution:

### 103. BLNN: Multiscale Feature Fusion-Based Bilinear Fine-Grained

摘要: Wood defects are quickly identified from an optical image based on deep learning methodology, which effectively improves the wood utilization. The traditional neural network technique is unemployed for the wood defect detection of optical image used, which results from a long training time, low recognition accuracy, and nonautomatic extraction of defect image features. In this paper, a wood knot defect detection model (so-called BLNN) combined deep learning is reported. Two subnetworks composed of convolutional neural networks are trained by Pytorch. By using the feature extraction capabilities of the two subnetworks and combining the bilinear join operation, the fine-grained features of the image are obtained. The experimental results show that the accuracy has reached up 99.20%, and the training time is obviously reduced with the speed of defect detection about 0.0795 s/image. It indicates that BLNN has the ability to improve the accuracy of defect recognition and has a potential application in the detection of wood knot defects.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
14  
ISSN:  
1687-725X  
eISSN:  
1687-7268  
入藏号:  
WOS:000691098400001  
Indexed Date:  
2021-09-07  
Institution Address:  
Northeast Forestry Univ, Coll Sci, Harbin 150040, Peoples R China  
Institution Address:  
Harbin Inst Technol, Sch Mechatron Engn, Harbin 150001, Peoples R China  
Institution Address:  
Harbin Inst Technol, State Key Lab Robot & Syst, Harbin 150001, Peoples R China  
Institution Address:  
Harbin Inst Technol, Sch Instrumentat Sci & Engn, Harbin 150001, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 104. FS-Net: Four-Stream Network With Spatial–Spectral Representation

摘要: In the field of hyperspectral and multispectral (HS-MS) image fusion, convolutional neural network (CNN) has achieved promising performance. Nevertheless, most networks cannot fully extract the spatial-spectral features from 3-D HS-MS images via vanilla convolution. 3-D-CNN, treating the spatial and spectral dimension at the same weight maybe a solution to this problem. However, this approach ignores the distinctiveness of spatial-spectral features and suffers from high computational complexity. To meet these challenges, a novel architecture called four-stream network (FS-Net) is proposed, which integrates the spatial-spectral feature extraction, and multiscale and multilevel feature fusion into a united framework. First, the simple but effective 2-D spatial feature extractor and spectral feature extractor are designed, which encode the spatial and spectral features of HS-MS images into four groups of feature maps. Next, these four groups of feature maps are fused in pairs, where the output can be seen as the comprehensive representation of input images. Finally, multilevel and multiscale feature fusion strategies helps to recover the HS image at high-spatial-resolution, and thus, simultaneously capture the detail texture and contextual information of the different and complex ground objects. Extensive experiments were conducted on three datasets with various experimental configurations to demonstrate that the proposed FS-Net achieves superior results quantitatively and visually when compared with the state-of-the-art networks.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001078150700015  
Indexed Date:  
2023-10-18  
Institution Address:  
Wuhan Univ, Sch Elect Informat, Wuhan 430072, Peoples R China  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & Re, Wuhan 430072, Peoples R China  
Institution:  
Institution:

### 105. Efficient Spectral-Spatial Fusion With Multiscale and Adaptive Attention

摘要: In hyperspectral image (HSI) classification, convolutional neural networks (CNNs) are widely used due to their ability to leverage the rich spectral information across multiple bands. However, HSI classification still faces various challenges, including insufficient spectral-spatial representation, excessive redundant information, and difficulties in effectively integrating features of different scales, etc., which may lead to reduced classification accuracy. In order to reduce the computational cost and improve the classification accuracy of land cover categories, an efficient spectral-spatial fusion method (ESSF) is proposed, which is based on the following modules: a multiscale feature fusion module (MSFFM), an efficient adaptive spectral-spatial feature extraction module (EASSFEM), and a context-aware fusion network (CFN). First, the MSFFM utilizes CNNs to extract and fuse features from various scales to comprehensively capture detailed spectral information in HSIs. Second, the EASSFEM dynamically adjusts the feature extraction process to optimize the fusion and representation of spectral-spatial features. In addition, it incorporates an adaptive attention mechanism to enhance the focus on relevant spectral-spatial features. Finally, the CFN enhances the model's ability to understand contextual relationships within the images, thereby improving classification accuracy. Extensive experiments conducted on four public datasets (Houston2013, Botswana, WHU-Hi-HanChuan and WHU-Hi-HongHu) demonstrate that the proposed ESSF method significantly outperforms nine other state-of-the-art methods in terms of classification accuracy.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001373734500002  
Indexed Date:  
2024-12-19  
Institution Address:  
Hengyang Normal Univ, Coll Comp Sci & Technol, Hengyang 421002, Peoples R China  
Institution Address:  
Hunan Prov Key Lab Intelligent Informat Proc & App, Hengyang 421002, Peoples R China  
Institution:  
Institution:  
Hunan Prov Key Lab Intelligent Informat Proc & App

### 106. MFI-Net: Multiscale Feature Interaction Network for Retinal Vessel

摘要: Segmentation of retinal vessels on fundus images plays a critical role in the diagnosis of micro-vascular and ophthalmological diseases. Although being extensively studied, this task remains challenging due to many factors including the highly variable vessel width and poor vessel-background contrast. In this paper, we propose a multiscale feature interaction network (MFI-Net) for retinal vessel segmentation, which is a U-shaped convolutional neural network equipped with the pyramid squeeze-and-excitation (PSE) module, coarse-to-fine (C2F) module, deep supervision, and feature fusion. We extend the SE operator to multiscale features, resulting in the PSE module, which uses the channel attention learned at multiple scales to enhance multiscale features and enables the network to handle the vessels with variable width. We further design the C2F module to generate and re-process the residual feature maps, aiming to preserve more vessel details during the decoding process. The proposed MFI-Net has been evaluated against several public models on the DRIVE, STARE, CHASE\_DB1, and HRF datasets. Our results suggest that both PSE and C2F modules are effective in improving the accuracy of MFI-Net, and also indicate that our model has superior segmentation performance and generalization ability over existing models on four public datasets.  
在 BIOSIS Citation Index 中的被引频次:  
9  
Web of Science 核心合集中的 "被引频次":  
35  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
38  
ISSN:  
2168-2194  
eISSN:  
2168-2208  
入藏号:  
WOS:000852247000022  
PubMed ID:  
35696471  
Indexed Date:  
2022-09-17  
Institution Address:  
Northwestern Polytech Univ, Natl Engn Lab Integrated Aerospace Ground Ocean B, Sch Comp Sci & Engn, Xian 710072, Shaanxi, Peoples R China  
Institution Address:  
Beihang Univ, Inst Artificial Intelligence, Beijing 100191, Peoples R China  
Institution Address:  
Monash Univ, Fac Informat Technol, Clayton, Vic 3800, Australia  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Informat & Commun Engn, Beijing 100876, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 107. Research on Defect Detection for Overhead Transmission Lines Based on

摘要: In the field of smart grid monitoring, real-time defect detection for overhead transmission lines is crucial for ensuring the safety and stability of power systems. This paper proposes a defect detection model for overhead transmission lines based on an improved YOLOv8n model, named ABG-YOLOv8n. The model incorporates four key improvements: Lightweight convolutional neural networks and spatial-channel reconstructed convolutional modules are integrated into the backbone network and feature fusion network, respectively. A bidirectional feature pyramid network is employed to achieve multi-scale feature fusion, and the ASFF mechanism is used to enhance the sensitivity of YOLOv8n's detection head. Finally, comprehensive comparative experiments were conducted with multiple models to validate the effectiveness of the proposed method based on the obtained prediction curves and various performance metrics. The validation results indicate that the proposed ABG-YOLOv8n model achieves a 4.5% improvement in mean average precision compared to the original YOLOv8n model, with corresponding increases of 3.6% in accuracy and 2.0% in recall. Additionally, the ABG-YOLOv8n model demonstrates superior detection performance compared to other enhanced YOLO models.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1996-1073  
入藏号:  
WOS:001377794400001  
Indexed Date:  
2024-12-21  
Institution Address:  
Shanghai Dianji Univ, Sch Elect Engn, Shanghai 201306, Peoples R China  
Institution:

### 108. Multi-Scale Feature Fusion Convolutional Neural Networks for Fault

摘要: Airborne electromechanical actuators (EMAs) play a key role in the flight control system, and their health condition has a considerable impact on the flight status and safety of aircraft. Considering the multi-scale feature of fault signals and the fault diagnosis reliability for EMAs under complex working conditions, a novel fault diagnosis method of multi-scale feature fusion convolutional neural network (MSFFCNN) is proposed. Leveraging the multiple different scales' learning structure and attention mechanism-based feature fusion, the fault-related information can be effectively captured and learned, thereby improving the recognition ability and diagnostic performance of the network. The proposed method was evaluated by experiments and compared with the other three fault-diagnosis algorithms. The results show that the proposed MSFFCNN approach has a better diagnostic performance compared with the state-of-the-art fault diagnosis methods, which demonstrates the effectiveness and superiority of the proposed method.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
7  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
eISSN:  
2076-3417  
入藏号:  
WOS:001045419300001  
Indexed Date:  
2023-08-19  
Institution Address:  
Xi An Jiao Tong Univ, State Key Lab Elect Insulat & Power Equipment, Xian 710049, Peoples R China  
Institution Address:  
State Grid Jibei Elect Power Co Ltd, Langfang Power Supply Co, Langfang 065000, Peoples R China  
Institution:  
Institution:

### 109. GAF-Net: A new automated segmentation method based on multiscale feature

摘要: Surface defect detection (SDD) is the necessary technique to monitor the surface quality of production. However, fine grain defects caused by stress loading, environmental influences, and construction defects is still a challenge to detect. In this research, the convolutional neural network for crack segmentation is developed based on the feature fusion and feedback on the global features and multi-scale feature (GAF-Net). First, a multi-scale feature feedback module (MSFF) is proposed, which uses four different scales to refine local features by fusing high-level and sub-high-level features to perform feedback correction. Secondly, the global feature module (GF) is proposed to generate a fine global information map using local features and adaptive weighted fusion with the correction map for crack detection. Finally, the GAF-Net network with multi-level feature maps is deeply supervised to accelerate GAF-Net and improve the detection accuracy. GAF-Net is trained and experimented on three publicly available pavement crack datasets, and the results show that GAF-Net achieves state-of-the-art results in the IoU segmentation metrics when compared to other deep learning methods (Crackforest: 53.61 %; Crack500: 65.19 %; DeepCrack: 81.63 %).  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
0167-8655  
eISSN:  
1872-7344  
入藏号:  
WOS:001370786100001  
Indexed Date:  
2024-12-13  
Institution Address:  
China Univ Geosci, Sch Mech Engn & Elect Informat, 388 LuMo Rd, Wuhan 430074, Peoples R China  
Institution Address:  
Chongqing Univ, State Key Lab Mech Transmiss Adv Equipment, Chongqing, Peoples R China  
Institution Address:  
China Univ Geosci, Shenzhen Res Inst, Shenzhen 518057, Peoples R China  
Institution:  
Institution:  
Institution:

### 110. A Novel 2D-3D CNN with Spectral-Spatial Multi-Scale Feature Fusion for

摘要: Multifarious hyperspectral image (HSI) classification methods based on convolutional neural networks (CNN) have been gradually proposed and achieve a promising classification performance. However, hyperspectral image classification still suffers from various challenges, including abundant redundant information, insufficient spectral-spatial representation, irregular class distribution, and so forth. To address these issues, we propose a novel 2D-3D CNN with spectral-spatial multi-scale feature fusion for hyperspectral image classification, which consists of two feature extraction streams, a feature fusion module as well as a classification scheme. First, we employ two diverse backbone modules for feature representation, that is, the spectral feature and the spatial feature extraction streams. The former utilizes a hierarchical feature extraction module to capture multi-scale spectral features, while the latter extracts multi-stage spatial features by introducing a multi-level fusion structure. With these network units, the category attribute information of HSI can be fully excavated. Then, to output more complete and robust information for classification, a multi-scale spectral-spatial-semantic feature fusion module is presented based on a Decomposition-Reconstruction structure. Last of all, we innovate a classification scheme to lift the classification accuracy. Experimental results on three public datasets demonstrate that the proposed method outperforms the state-of-the-art methods.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
20  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
22  
eISSN:  
2072-4292  
入藏号:  
WOS:000725527200001  
Indexed Date:  
2021-12-10  
Institution Address:  
Chinese Acad Sci, Changchun Inst Optic Fine Mech & Phys, Changchun 130033, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 111. An appearance quality classification method for <i>Auricularia

摘要: The intelligent appearance quality classification method for Auricularia auricula is of great significance to promote this industry. This paper proposes an appearance quality classification method for Auricularia auricula based on the improved Faster Region-based Convolutional Neural Networks (improved Faster RCNN) framework. The original Faster RCNN is improved by establishing a multiscale feature fusion detection model to improve the accuracy and real-time performance of the model. The multiscale feature fusion detection model makes full use of shallow feature information to complete target detection. It fuses shallow features with rich detailed information with deep features rich in strong semantic information. Since the fusion algorithm directly uses the existing information of the feature extraction network, there is no additional calculation. The fused features contain more original detailed feature information. Therefore, the improved Faster RCNN can improve the final detection rate without sacrificing speed. By comparing with the original Faster RCNN model, the mean average precision (mAP) of the improved Faster RCNN is increased by 2.13%. The average precision (AP) of the first-level Auricularia auricula is almost unchanged at a high level. The AP of the second-level Auricularia auricula is increased by nearly 5%. And the third-level Auricularia auricula AP is increased by 1%. The improved Faster RCNN improves the frames per second from 6.81 of the original Faster RCNN to 13.5. Meanwhile, the influence of complex environment and image resolution on the Auricularia auricula detection is explored.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2045-2322  
入藏号:  
WOS:001263443800051  
PubMed ID:  
38969651  
Indexed Date:  
2024-07-22  
Institution Address:  
Tianjin Agr Univ, Coll Engn & Technol, Tianjin 300392, Peoples R China  
Institution Address:  
Minist Agr & Rural Affairs TJAU, Key Lab Smart Breeding Coconstruct Minist & Prov, Tianjin 300392, Peoples R China  
Institution Address:  
Jiamusi Univ, Coll Mech Engn, Jiamusi 154007, Peoples R China  
Institution:  
Institution:  
Minist Agr & Rural Affairs TJAU  
Institution:

### 112. AFL-Net: Attentional Feature Learning Network for Building Extraction

摘要: Convolutional neural networks (CNNs) perform well in tasks of segmenting buildings from remote sensing images. However, the intraclass heterogeneity of buildings is high in images, while the interclass homogeneity between buildings and other nonbuilding objects is low. This leads to an inaccurate distinction between buildings and complex backgrounds. To overcome this challenge, we propose an Attentional Feature Learning Network (AFL-Net) that can accurately extract buildings from remote sensing images. We designed an attentional multiscale feature fusion (AMFF) module and a shape feature refinement (SFR) module to improve building recognition accuracy in complex environments. The AMFF module adaptively adjusts the weights of multi-scale features through the attention mechanism, which enhances the global perception and ensures the integrity of building segmentation results. The SFR module captures the shape features of the buildings, which enhances the network capability for identifying the area between building edges and surrounding nonbuilding objects and reduces the over-segmentation of buildings. An ablation study was conducted with both qualitative and quantitative analyses, verifying the effectiveness of the AMFF and SFR modules. The proposed AFL-Net achieved 91.37, 82.10, 73.27, and 79.81% intersection over union (IoU) values on the WHU Building Aerial Imagery, Inria Aerial Image Labeling, Massachusetts Buildings, and Building Instances of Typical Cities in China datasets, respectively. Thus, the AFL-Net offers the prospect of application for successful extraction of buildings from remote sensing images.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
6  
被引频次合计:  
8  
eISSN:  
2072-4292  
入藏号:  
WOS:000909210000001  
Indexed Date:  
2023-01-24  
Institution Address:  
PLA Strateg Support Force Informat Engn Univ, Inst Geospatial Informat, Zhengzhou 450001, Peoples R China  
Institution:

### 113. Real-Time Quality Inspection of Motor Rotor Using Cost-Effective

摘要: Induction motors (IMs) are used extensively as driving actuators in electric vehicles. Motor rotors are prone to defects in the die casting procedure, which can significantly reduce the production quality. Benefitting from the development of Internet of Things (IoT) techniques and edge computing, this study designed an instrumentation system for the fast inspection of rotor defects to meet the objectives of efficient and high-quality rotor production. First, an electromagnetic sensing device is designed to acquire the induced voltage signal of the rotor under investigation. Second, a residual multiscale feature fusion convolutional neural network model is designed to extract the hierarchical features of the signal, to facilitate defect recognition. The developed algorithm is deployed into a cost-effective edge computing node that includes a signal acquisition circuit and a Raspberry Pi microcontroller. The conducted experimental studies show that this implementation can achieve an inference time of less than 200 ms and accuracy of more than 99%. It is shown that the designed system exhibits superior performance when compared with conventional methods. The developed, compact and flexible handheld solution with enhanced deep learning techniques shows outstanding potential for use in real-time rotor defect detection.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
13  
被引频次合计:  
13  
ISSN:  
2327-4662  
入藏号:  
WOS:000968830500067  
Indexed Date:  
2023-05-14  
Institution Address:  
Anhui Univ, Coll Elect Engn & Automat, Hefei 230601, Peoples R China  
Institution Address:  
Anhui Univ, Coll Elect & Informat Engn, Hefei 230601, Peoples R China  
Institution Address:  
Univ Sci & Technol China, Dept Precis Machinery & Precis Instrumentat, Hefei 230027, Peoples R China  
Institution Address:  
Univ British Columbia, Dept Mech Engn, Vancouver, BC V6T 1Z4, Canada  
Institution Address:  
Univ Lancaster, Dept Engn, Lancaster LA1 4YW, England  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 114. MRFT: Multiscale Recurrent Fusion Transformer Based Prior Knowledge for

摘要: Bit-depth enhancement (BDE) plays an important role in providing high bit-depth data support for high-dynamic range (HDR) display. Although convolutional neural network (CNN) based BDE methods have achieved top performance, multiscale feature extraction and fusion still suffer from some inherent architectural flaws. Moreover, the training-data-scarce scene has not been effectively explored. To this end, this paper proposes an innovative multiscale recurrent fusion transformer (MRFT) framework, which contains three key components, i.e. multiscale transformer feature encoder, recurrent feature fusion module, and prior knowledge injection. Specifically, the multiscale transformer feature encoder consists of a prior-injected context encoder (PICE) and a multiscale local feature encoder (MLFE). PICE leverages the vanilla self-attention mechanism to extract the global context correlating spatially-distant contents for distinguishing long-distance false contours. MLFE exploits the local self-attention mechanism with varied window sizes to capture different-scale detail features. Then, a hierarchical recurrent decoder (HRD) is proposed as the recurrent feature fusion module to fuse multiscale visual information with global guidance. Via the circular query-key mechanism, global-to-local information is progressively fused. Furthermore, we propose a two-stage alternating optimization strategy for prior knowledge injection. By pre-parameterizing the global auxiliary priors, the training dilemma on the data-scarce domain is significantly alleviated. Extensive analyses on multiple benchmark datasets demonstrate the superiority of our MRFT in terms of quantitative measures and aesthetic effects.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
4  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:001082387900015  
Indexed Date:  
2023-11-09  
Institution Address:  
Tianjin Univ, Sch Elect & Informat Engn, Tianjin 300072, Peoples R China  
Institution:

### 115. A Parallel Image Denoising Network Based on Nonparametric Attention and

摘要: Convolutional neural networks have achieved excellent results in image denoising; however, there are still some problems: (1) The majority of single-branch models cannot fully exploit the image features and often suffer from the loss of information. (2) Most of the deep CNNs have inadequate edge feature extraction and saturated performance problems. To solve these problems, this paper proposes a two-branch convolutional image denoising network based on nonparametric attention and multiscale feature fusion, aiming to improve the denoising performance while better recovering the image edge and texture information. Firstly, ordinary convolutional layers were used to extract shallow features of noise in the image. Then, a combination of two-branch networks with different and complementary structures was used to extract deep features from the noise information in the image to solve the problem of insufficient feature extraction by the single-branch network model. The upper branch network used densely connected blocks to extract local features of the noise in the image. The lower branch network used multiple dilation convolution residual blocks with different dilation rates to increase the receptive field and extend more contextual information to obtain the global features of the noise in the image. It not only solved the problem of insufficient edge feature extraction but also solved the problem of the saturation of deep CNN performance. In this paper, a nonparametric attention mechanism is introduced in the two-branch feature extraction module, which enabled the network to pay attention to and learn the key information in the feature map, and improved the learning performance of the network. The enhanced features were then processed through the multiscale feature fusion module to obtain multiscale image feature information at different depths to obtain more robust fused features. Finally, the shallow features and deep features were summed using a long jump join and were processed through an ordinary convolutional layer and output to obtain a residual image. In this paper, Set12, BSD68, Set5, CBSD68, and SIDD are used as a test dataset to which different intensities of Gaussian white noise were added for testing and compared with several mainstream denoising methods currently available. The experimental results showed that this paper's algorithm had better objective indexes on all test sets and outperformed the comparison algorithms. The method in this paper not only achieved a good denoising effect but also effectively retained the edge and texture information of the original image. The proposed method provided a new idea for the study of deep neural networks in the field of image denoising.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001405283900001  
PubMed ID:  
39860687  
Indexed Date:  
2025-02-01  
Institution Address:  
Univ Kitakyushu, Grad Sch Environm Engn, Kitakyushu 8080135, Japan  
Institution Address:  
Univ Kitakyushu, Dept Informat & Syst Engn, Kitakyushu 8080135, Japan  
Institution Address:  
Ankang Univ, Sch Elect & Informat Engn, Ankang 725000, Peoples R China  
Institution:  
Institution:  
Institution:

### 116. Hybrid CNN and Transformer Network for Semantic Segmentation of UAV

摘要: Semantic segmentation of unmanned aerial vehicle (UAV) remote sensing images is a recent research hotspot, offering technical support for diverse types of UAV remote sensing missions. However, unlike general scene images, UAV remote sensing images present inherent challenges. These challenges include the complexity of backgrounds, substantial variations in target scales, and dense arrangements of small targets, which severely hinder the accuracy of semantic segmentation. To address these issues, we propose a convolutional neural network (CNN) and transformer hybrid network for semantic segmentation of UAV remote sensing images. The proposed network follows an encoder-decoder architecture that merges a transformer-based encoder with a CNN-based decoder. First, we incorporate the Swin transformer as the encoder to address the limitations of CNN in global modeling, mitigating the interference caused by complex background information. Second, to effectively handle the significant changes in target scales, we design the multiscale feature integration module (MFIM) that enhances the multiscale feature representation capability of the network. Finally, the semantic feature fusion module (SFFM) is designed to filter the redundant noise during the feature fusion process, which improves the recognition of small targets and edges. Experimental results demonstrate that the proposed method outperforms other popular methods on the UAVid and Aeroscapes datasets.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
2576-3164  
入藏号:  
WOS:001388140300007  
Indexed Date:  
2025-01-09  
Institution Address:  
Northeast Petr Univ, Sch Comp & Informat Technol, Daqing 163318, Peoples R China  
Institution Address:  
Changshu Inst Technol, Sch Comp Sci & Engn, Suzhou 215500, Peoples R China  
Institution Address:  
China Meteorol Adm, Natl Satellite Meteorol Ctr, Key Lab Radiometr Calibrat & Validat Environm Sate, Beijing 100080, Peoples R China  
Institution Address:  
Changshu Inst Technol, Sch Elect & Automat Engn, Suzhou 215500, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 117. MS3Net: Multiscale stratified-split symmetric network with quadra-view

摘要: Recently, hyperspectral image (HSI) classification has become a promising research direction in remote sensing image processing. Many HSI classification methods have been proposed based on convolutional neural networks (CNNs) and attention mechanisms (AMs). However, most current CNN-based methods only consider extracting features at a single scale in HSI, which may ignore the delicate features of some objects. Moreover, present AMs primarily focus on one feature dimension, such as spatial or channel attention, while disregarding dimension interaction. To conquer the above issues, a novel multiscale stratified-split symmetric network with quadra-view attention, namely MS3Net, is proposed for HSI classification. Generally, the proposed MS3Net has a dual-stream symmetric pipeline, which can better extract HSI's spectral signatures and spatial features. Specifically, the proposed MS3Net consists of three modules: a multiscale feature extraction module, a feature enhancement module, and a feature fusion module. Firstly, a stratified-split module is designed to extract multiscale spectral and spatial features. In addition, to reduce the complexity of the model, we designed pseudo-3-D spectral and spatial convolution to replace the traditional 3-D convolution operation. Secondly, a novel quadra-view attention module is proposed, guiding the model to focus on important features from multiple dimensions. Finally, the selective kernel feature fusion module is introduced, which can dynamically integrate spectral and spatial features. Experimental results on four benchmark HSI datasets with different scenes and resolutions confirm the visual and quantitative superiority of the proposed MS3Net over the state-of-the-art related methods in this research direction.  
Web of Science 核心合集中的 "被引频次":  
8  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
0165-1684  
eISSN:  
1872-7557  
入藏号:  
WOS:001035433700001  
Indexed Date:  
2023-08-10  
Institution Address:  
Qiqihar Univ, Coll Comp & Control Engn, Qiqihar 161006, Peoples R China  
Institution Address:  
Dalian Minzu Univ, Coll Informat & Commun Engn, Dalian 116600, Peoples R China  
Institution:  
Institution:

### 118. Adapting Segment Anything Model to Aerial Land Cover Classification With

摘要: Recently, vision foundation models have gathered wide attention. Among the many endeavors, the segment anything model (SAM) makes remarkable progress toward a universal model showing unprecedented generalization ability. We propose a novel semantic segmentation model that combines SAM's image encoder and a low-rank adaptation (LoRA) approach for feature extraction and fine-tuning on aerial images. We also employ an auxiliary CNN encoder to facilitate downstream adaptation and complement the ViT encoder on dense vision tasks. Furthermore, cross-attention is utilized to implement feature interactions between the two encoders. Finally, the UperNet head is employed for multiscale feature fusion and generating segmentation masks. The proposed model was evaluated on the ISPRS Vaihingen and Potsdam datasets and achieved the best mean intersection-over-union (mIoU) of 76.44 and 78.01 for the two datasets. The evaluation results demonstrate the superiority of our model.  
Web of Science 核心合集中的 "被引频次":  
16  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
16  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001230653000001  
Indexed Date:  
2024-05-31  
Institution Address:  
Geovis Spatial Technol Co Ltd, Xian 710054, Peoples R China  
Institution Address:  
Harbin Inst Technol Shenzhen, Inst Space Sci & Appl Technol, Shenzhen 518057, Peoples R China  
Institution Address:  
Geovis Spatial Technol Co Ltd, Xian 710054, Peoples R China  
Institution Address:  
Changan Univ, Coll Geol Engn & Geomat, Xian 710064, Peoples R China  
Institution:  
Geovis Spatial Technol Co Ltd  
Institution:  
Institution:  
Geovis Spatial Technol Co Ltd  
Institution:

### 119. A Multi-Task-Based Deep Multi-Scale Information Fusion Method for

摘要: The use of deep learning for fault diagnosis is already a common approach. However, integrating discriminative information of fault types and scales into deep learning models for rich multitask fault feature diagnosis still deserves attention. In this study, a deep multitask-based multiscale feature fusion network model (MEAT) is proposed to address the limitations and poor adaptability of traditional convolutional neural network models for complex jobs. The model performed multidimensional feature extraction through convolution at different scales to obtain different levels of fault information, used a hierarchical attention mechanism to weight the fusion of features to achieve an accuracy of 99.95% for the total task of fault six classification, and considered two subtasks in fault classification to discriminate fault size and fault type through multi-task mapping decomposition. Of these, the highest accuracy of fault size classification reached 100%. In addition, Precision, ReCall, and Sacore F1 all reached the index of 1, which achieved the accurate diagnosis of bearing faults.  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
eISSN:  
2075-1702  
入藏号:  
WOS:000941545800001  
Indexed Date:  
2023-03-18  
Institution Address:  
Jilin Inst Chem Technol, Sch Informat & Control Engn, Jilin 132022, Peoples R China  
Institution Address:  
Jilin Inst Chem Technol, Sch Sci, Jilin 132022, Peoples R China  
Institution Address:  
Jilin Univ, Sch Software, Changchun 130015, Peoples R China  
Institution:  
Institution:  
Institution:

### 120. ORCNN-X: Attention-Driven Multiscale Network for Detecting Small Objects

摘要: Currently, object detection on remote sensing images has drawn significant attention due to its extensive applications, including environmental monitoring, urban planning, and disaster assessment. However, detecting objects in the aerial images captured by remote sensors presents unique challenges compared to natural images, such as low resolution, complex backgrounds, and variations in scale and angle. Prior object detection algorithms are limited in their ability to identify oriented small objects, especially in aerial images where small objects are usually obscured by background noise. To address the above limitations, a novel framework (ORCNN-X) was proposed for oriented small object detection in remote sensing images by improving the Oriented RCNN. The framework adopts a multiscale feature extraction network (ResNeSt+) with a dynamic attention module (DCSA) and an effective feature fusion mechanism (W-PAFPN) to enhance the model's perception ability and handle variations in scale and angle. The proposed framework is evaluated based on two public benchmark datasets, DOTA and HRSC2016. The experiments demonstrate its state-of-the-art performance in aspects of detection accuracy and speed. The presented model can also represent more objective spatial location information according to the feature visualization maps. Specifically, our model outperforms the baseline model by 1.43% mAP50 and 1.37% mAP(12) on DOTA and HRSC2016 datasets, respectively.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2072-4292  
入藏号:  
WOS:001038916100001  
Indexed Date:  
2023-08-10  
Institution Address:  
Qufu Normal Univ, Sch Comp Sci, Rizhao 276826, Peoples R China  
Institution Address:  
Sejong Univ, Dept Informat & Commun Engn, Seoul 05006, South Korea  
Institution Address:  
Sejong Univ, Convergence Engn Intelligent Drone, Seoul 05006, South Korea  
Institution Address:  
Sejong Univ, Dept Comp Sci & Engn, Seoul 05006, South Korea  
Institution:  
Institution:  
Institution:  
Institution:

### 121. Exploiting multi-scale hierarchical feature representation for visual

摘要: Convolutional neural networks (CNNs) have been the dominant architectures for feature extraction tasks, but CNNs do not look for and focus on some specific image features. Correlation operations play an important role in visual tracking. However, the correlation operation reserves a large amount of unfavorable background information. In this paper, we propose an effective feature recognizer including channel and spatial attention modules to focus on important object feature information. Thus, the representation power of the feature extraction network is improved. Further, we design a multi-scale feature fusion network. The fusion network performs feature fusion on template feature and encoded feature branches to establish connections between features at different scales. Experiments on six benchmarks demonstrate that the proposed tracker outperforms the state-of-the-art trackers. In particular, the proposed tracker achieves an 80.4% AUC on TrackingNet and a 68.4% AUC on GOT-10k while running at a real-time speed.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2199-4536  
eISSN:  
2198-6053  
入藏号:  
WOS:001160577300002  
Indexed Date:  
2024-02-20  
Institution Address:  
Nanchang Inst Technol, Sch Informat Engn, Nanchang 330029, Peoples R China  
Institution Address:  
Nanchang Inst Technol, Jiangxi Prov Key Lab Water Informat Cooperat Sensi, Nanchang 330029, Peoples R China  
Institution:  
Institution:

### 122. A convolutional neural network model of multi-scale feature fusion:

摘要: MFF-Net (a multi-scale feature fusion convolutional neural network) was designed to improve the recognition rate of handwritten digits. The low-level, middle-level and high-level features of the image were first extracted through the convolution operation, and then the low-level and intermediate features were further extracted through different convolutional layers, later directly fused with the high-level features of the image with a certain weight, and then processed by the full connection layer. By adding a batch normalization layer before the activation layer, and a dropout layer between the full connection layers, the accuracy and generalization capacity of the network are improved. At the same time, a dynamic learning rate algorithm was designed, with which, the trained network accuracy was significantly improved as shown in the experiments on the MNIST data set. The accurate rate could reach 99.66% through only 30 epochs training. The comparison indicated that the accuracy of the network model is significantly higher than that of others.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
1472-7978  
eISSN:  
1875-8983  
入藏号:  
WOS:000893587400029  
Indexed Date:  
2022-12-28  
Institution Address:  
Anhui Polytech Univ, Sch Comp & Informat, Wuhu, Anhui, Peoples R China  
Institution Address:  
Anhui Normal Univ, Sch Comp & Informat, Wuhu 241000, Anhui, Peoples R China  
Institution:  
Institution:

### 123. Adaptive multi-scale feature channel grouping optimization algorithm

摘要: Aiming at the balance optimization problem of Lightweight Convolutional Neural Network (LCNN) in accuracy and complexity, an adaptive multi-scale feature channel grouping optimization algorithm based on fast Nondominated Sorting Genetic Algorithm (NSGA-Ⅱ) was proposed to optimize the feature channel grouping structure of LCNN. Firstly, the complexity minimization and accuracy maximization of the feature fusion layer structure in LCNN were regarded as two optimization objectives, and the dual-objective function modeling and theoretical analysis were carried out. Then, a LCNN structure optimization framework based on NSGA-Ⅱ was designed, and an adaptive grouping layer based on NSGA-Ⅱ was added to deep convolution layer in original LCNN structure, thus constructing an Adaptive Multi-scale Feature Fusion Network based on NSGA2 (NSGA2-AMFFNetwork). Experimental results on image classification datasets show that compared with the manually designed network structure M\_blockNet\_v1, NSGA2-AMFFNetwork has the average accuracy improved by 1.220 2 percentage points, and the running time decreased by 41.07%. This above indicates that the proposed optimization algorithm can balance the complexity and accuracy of LCNN, and also provide more options for network structure with balanced performance for ordinary users who lack domain knowledge.  
摘要:  
针对轻量型卷积神经网络(LCNN)的精确度和复杂度均衡优化问题,提出基于快速非支配排序遗传算法(NSGA-Ⅱ)的自适应多尺度特征通道分组优化算法对LCNN特征通道分组结构进行优化。首先,将LCNN中的特征融合层结构的复杂度最小化和精确度最大化作为两个优化目标,进行双目标函数建模及理论分析;然后,设计基于NSGA-Ⅱ的LCNN结构优化框架,并在原始LCNN结构的深度卷积层之上增加基于NSGA-Ⅱ的自适应分组层,构建基于NSGA-Ⅱ的自适应多尺度的特征融合网络NSGA2-AMFFNetwork。在图像分类数据集上的实验结果显示,与手工设计的网络结构M\_blockNet\_v1相比,NSGA2-AMFFNetwork的平均精确度提升了1.220 2个百分点,运行时间降低了41.07%。这表明所提优化算法能较好平衡LCNN的复杂度和精确度,同时还可为领域知识不足的普通用户提供更多性能表现均衡的网络结构选择方案。  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1001-9081  
入藏号:  
CSCD:7463997  
Indexed Date:  
2023-08-18  
Institution Address:  
西安理工大学计算机科学与工程学院;;西安理工大学, ;;陕西省网络计算与安全技术重点实验室, 西安;;西安, ;; 710048;;710048  
Institution Address:  
西安理工大学计算机科学与工程学院, 西安, 陕西 710048, 中国  
Institution Address:  
School of Computer Science and Engineering, Xi'an University of Technology;;Xi'an University of Technology, ;;Shaanxi Key Laboratory for Network Computing and Security Technology, Xi'an;;Xi'an, Shaanxi;;Shaanxi 710048;;710048  
Institution Address:  
School of Computer Science and Engineering, Xi'an University of Technology, Xi'an, Shaanxi 710048, China  
Institution:  
西安理工大学计算机科学与工程学院;;西安理工大学  
Institution:  
西安理工大学计算机科学与工程学院  
Institution:  
School of Computer Science and Engineering, Xi'an University of Technology;;Xi'an University of Technology  
Institution:  
School of Computer Science and Engineering, Xi'an University of Technology

### 124. Monocular depth estimation with enhanced edge

摘要: To solve the problem that the discrimination of object occlusion is difficult and the accuracy of depth estimation at the boundary is low caused by unobvious depth jump in monocular depth estimation,a method of monocular depth estimation with enhanced edge was proposed.Two sets of convolutional neural networks were applied to the algorithm,depth estimation network output initially predicted depth map and depth compensation network output compensate value for depth prediction,and the output of two sets of networks were fused to achieve the depth compensation of the original prediction around the object boundary.The multiscale feature fusion loss function was introduced,and the loss function of point constraint was designed to further improve the depth estimation accuracy at the boundary. Through the test and verification on NYU Depth v2 dataset and iBims dataset,the experimental results show that the method in this paper can effectively improve the clarity of the object profile in the depth map, make it easier for the object to block and judge and further enhance the effect of monocular depth estimation.  
摘要:  
为解决目前单目图像深度估计过程中物体边界处深度跳变不明显导致的遮挡难以判别、边界处深度估计准确度较低的问题,提出了一种强化边缘的单目图像深度估计方法.采用深度估计网络输出最初预测的深度图,同时采用深度补偿网络输出应补偿深度的预测值,通过融合两组网络的输出实现对最初预测的深度图中物体边界轮廓处深度值的补偿.此外,通过设计点约束损失函数,并引入多尺度特征融合损失函数进一步提升边界处的深度估计精度.在NYU Depth v2数据集和iBims数据集上的测试实验表明本文方法能有效提升深度图中物体轮廓的清晰度,使得物体遮挡判别更加容易,可进一步提升单目图像深度估计的效果.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1671-4512  
入藏号:  
CSCD:7183653  
Indexed Date:  
2022-06-17  
Institution Address:  
武汉大学电子信息学院, 武汉, 湖北 430072, 中国  
Institution Address:  
School of Electrical Information,Wuhan University, Wuhan, Hubei 430072, China  
Institution:  
武汉大学电子信息学院  
Institution:  
School of Electrical Information,Wuhan University

### 125. MSFMamba: Multiscale Feature Fusion State Space Model for Multisource

摘要: In the field of multisource remote sensing image classification, remarkable progress has been made by using the convolutional neural network (CNN) and Transformer. While CNNs are constrained by their local receptive fields, Transformers mitigate this issue with their global attention mechanism. However, Transformers come with the tradeoff of higher computational complexity. Recently, Mamba-based methods built upon the state space model (SSM) have shown great potential for long-range dependence modeling with linear complexity, but they have rarely been explored for multisource remote sensing image classification tasks. To address this issue, we propose the Multi-Scale Feature Fusion Mamba (MSFMamba) network, a novel framework designed for the joint classification of hyperspectral image (HSI) and light detection and ranging (LiDAR)/synthetic aperture radar (SAR) data. The MSFMamba network is composed of three key components: the Multi-Scale Spatial Mamba (MSpa-Mamba) block, the Spectral Mamba (Spe-Mamba) block, and the fusion Mamba (Fus-Mamba) block. The MSpa-Mamba block employs a multiscale strategy to reduce computational cost and alleviate feature redundancy in multiple scanning routes, ensuring efficient spatial feature modeling. The Spe-Mamba block focuses on spectral feature extraction, addressing the unique challenges of HSI data representation. Finally, the Fus-Mamba block bridges the heterogeneous gap between HSI and LiDAR/SAR data by extending the original Mamba architecture to accommodate dual inputs, enhancing cross-modal feature interactions and enabling seamless data fusion. Together, these components enable MSFMamba to effectively tackle the challenges of multisource data classification, delivering improved performance with optimized computational efficiency. Comprehensive experiments on four real-world multisource remote sensing datasets (Berlin, Augsburg, Houston2018, and Houston2013) demonstrate the superiority of MSFMamba outperforms several state-of-the-art methods and achieves overall accuracies of 76.92%, 91.38%, 92.38%, and 92.86%, respectively. The source codes of MSFMamba will be publicly available at https://github.com/oucailab/MSFMamba.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001416230900023  
Indexed Date:  
2025-02-14  
Institution Address:  
Ocean Univ China, Sch Comp Sci & Technol, Qingdao 266100, Peoples R China  
Institution Address:  
Mississippi State Univ, Dept Elect & Comp Engn, Starkville, MS 39762 USA  
Institution:  
Institution:

### 126. Similar-Signal Recognition Method for phi-OTDR Systems Based on

摘要: Objective A phase-sensitive optical time-domain reflectometer (phi-OTDR) system is a front monitoring and early warning technology that can acquire the location of disturbances in space and phase information of disturbances in time.With the advantages of high resolution,wide monitoring range,and strong anti-interference capability,this technology has been widely used in pipeline safety maintenance,intrusion warning,and large-equipment monitoring.However,due to the complex diversity of the application environment,the system suffers from low recognition accuracy and insufficient stability in actual use,particularly when similar signals are recognized in the system application.To solve these problems,this study proposes a similar-signal recognition method based on multiscale feature fusion.This method can effectively improve the recognition accuracy of similar signals while maintaining the recognition accuracy of the base signal.Methods The original signal is first decomposed into sub-signals in different frequency ranges using empirical mode decomposition(EMD) and wavelet packet decomposition (WPD).The original signal and individual sub-signals are then subjected to time-frequency feature extraction and approximate entropy feature extraction.The time-frequency features are used to evaluate the details of the time and frequency variations of the signal,the approximate entropy features are used to evaluate the complexity and regularity of the signal,and the multiscale signal decomposition and multi-feature extraction are used to amplify the feature differences between similar signals.Because the multiscale and multi-feature approach increases the dimensionality of the data,the proposed method utilizes principal component analysis (PCA) to combine high-dimensional features and reduce the dimensionality of system features,thereby improving system efficiency.Finally,the fused features are passed into a lightweight back-propagation (BP) neural network as input variables for signal data processing.Compared to other traditional neural networks,BP neural networks have the advantages of lightweight structures and high speed,enabling them to process signal data quickly.Results and Discussions Sub-signals decomposed by EMD and WPD have multiscale characteristics ranging from low to high frequencies.Each sub-signal contains a part of the signal domain within the main frequency-band range of the original data.Decomposition helps to amplify the feature gaps between different signals and facilitates subsequent multidimensional feature extraction (Fig.10).Following feature extraction and fusion,the four signals show significant differences in the feature space.Thus,even with a simple classifier,signal classification and recognition can be achieved (Fig.11).A comparison among extracting multi-features from original signal [Fig.12(a)],the CNN model [Fig.12(b)],and the multi-scale feature fusion[Fig.12(c)] reveals that the multi-scale feature fusion has higher recognition accuracy,where knocking and shakingsignal recognition accuracies reach 100% and trolleying and walking-signal recognition accuracies reach 98.5% and 98.0%,respectively.A comprehensive analysis reveals that the comprehensive recognition accuracy of the proposed method is increased by 8.4 and 9.0 percentage points over extracting multi-features from original signal and CNN model,respectively,and the similar-signal recognition accuracy is increased by 13.5 and 12.4 percentage points (Fig.13),respectively.These results verify that the method has high recognition accuracy.Conclusions Experimental results show that the decomposition method using EMD combined with WPD can obtain sub-signals at different scales.The time-frequency domain and approximate entropy features can in turn be extracted from the original signal and subsignal to enhance the differentiation of similar-signal features more effectively.  
摘要:  
为解决分布式相位敏感光时域反射计系统现有事件识别方法对于相似振动信号识别困难这一问题,提出了一种基于多尺度特征融合的相似信号识别方法。在该方法中,原始信号首先通过经验模态分解和小波包分解被分解为不同频率范围内的子信号。随后,分别提取原始信号和子信号的时频特征和近似熵特征,并利用主成分分析法对所提取的特征进行融合。最后,通过构建一个6层轻量反向传播(BP)神经网络分类器,训练分类模型并利用测试集验证模型分类度。该方法对小车经过和行走等相似信号的识别准确率可分别达到98.5%和98.0%,对于敲击和摇晃差异性大的信号的识别准确率可达100%。相比于直接从原始信号中提取特征并结合时频图的卷积神经网络方式,所提方法的综合识别准确率分别提高了8.4%与9.0%,相似信号的识别准确率分别提高了13.5%与12.4%。结果表明,该方法在保证差异性大的信号的高识别准确率的基础上,显著提高了相似信号的识别准确率,对于拓展分布式光纤传感的应用范围有重要的价值。  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0258-7025  
入藏号:  
CSCD:7706686  
Indexed Date:  
2024-07-06  
Institution Address:  
中国计量大学光学与电子科技学院, 杭州, 浙江 310018, 中国  
Institution Address:  
College of Optical and Electronic Technology,China Jiliang University, Hangzhou, Zhejiang 310018, China  
Institution:  
中国计量大学光学与电子科技学院  
Institution:  
College of Optical and Electronic Technology,China Jiliang University

### 127. A Multi-Scale Feature Fusion Convolutional Neural Network Approach for

摘要: A multi-scale feature fusion convolutional neural network-based SSVEP signal classification and recognition method is proposed to solve the problems of low classification accuracy,inadequate feature extraction,complex and time-consuming methods of traditional steady-state visual evoked potential(SSVEP-MF)signal target recognition methods.Firstly,the wavelet transform is used to integrate the multi-channel SSVEP signals into two-dimensional images as the input sample set;secondly,a multi-scale feature fusion convolutional neural network model(MFCNN)is established,which uses a three-layer two-dimensional convolutional kernel to achieve sufficient extraction of features at different scales of image samples,constructs multi-scale feature fusion units to fuse features at different levels,and completes the training of the model through operations such as full connectivity;finally,the sample set is input to the MFCNN model to achieve adaptive extraction of EEG signal features and end-to-end classification.The proposed SSVEP-MF method can fully extract the features at each level of the signal,achieve effective recognition of SSVEP signals under short-time visual stimulation,and have high target recognition efficiency.The experimental results show that the recognition accuracy of the proposed method is improved by 18.57%,20.08% and 7.03%,respectively, compared with the traditional power spectral density analysis method,typical correlation analysis method and common convolutional structure method at 1sstimulus duration,which effectively improves the signal recognition performance of brain-machine interface based on the steady-state visual evoked potential paradigm.  
摘要:  
针对传统稳态视觉诱发电位(SSVEP)脑电信号目标识别方法分类精度低、提取特征不充分、方法复杂且耗时等问题,提出一种基于多尺度特征融合卷积神经网络的SSVEP信号分类识别方法(SSVEP-MF)。利用小波变换将多通道SSVEP信号整合转化为二维图像作为输入样本集;建立多尺度特征融合卷积神经网络模型(MFCNN),该模型利用三层二维卷积核实现图像样本不同尺度特征的充分提取,构建多尺度特征融合单元对不同层级特征进行融合,并通过全连接等操作完成模型的训练;将样本集输入到MFCNN模型中实现脑电信号特征自适应提取及端到端分类。所提SSVEP-MF方法能够充分提取信号各层级特征,实现短时间视觉刺激下SSVEP信号的有效识别,并具有较高的目标识别效率。实验结果表明,在1s刺激时长时,相比传统功率谱密度分析方法、典型相关分析方法以及普通卷积结构方法,所提方法的识别准确率分别提升了18.57%、 20.08%及7.03%,有效提高了基于稳态视觉诱发电位范式下脑机接口的信号识别性能。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
1  
ISSN:  
0253-987X  
入藏号:  
CSCD:7194315  
Indexed Date:  
2022-07-01  
Institution Address:  
新疆大学机械工程学院, 乌鲁木齐, 新疆 830047, 中国  
Institution Address:  
西安交通大学机械工程学院, 西安, 陕西 710049, 中国  
Institution Address:  
School of Mechanical Engineering, Xinjiang University, Urumqi, Xinjiang 830047, China  
Institution Address:  
School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China  
Institution:  
新疆大学机械工程学院  
Institution:  
西安交通大学机械工程学院  
Institution:  
School of Mechanical Engineering, Xinjiang University  
Institution:  
School of Mechanical Engineering, Xi'an Jiaotong University

### 128. Facial Expression Recognition Network Based on Attention Mechanism

摘要: Facial expression recognition has remained a challenging problem in computer vision. Recently,with the rapid development of deep learning,some methods based on convolutional neural networks have greatly improved the accuracy of facial expression recognition. However,these methods have not fully used the available information because the meaningful features for facial expression recognition are mainly concentrated in some key locations,such as eyes,nose,and mouth. Increasing the weight of these key positions can improve the effect of facial expression recognition. This paper proposed a facial expression recognition network based on an attention mechanism. First,a deep and shallow feature fusion structure was added to the backbone network. This structure was designed to fully extract the shallow features at various scales from the original image and cascade these features with deep features to reduce information loss during forward propagation. Second,a two-step-based channel attention module was embedded in the network to encode the channel information in the cascaded feature map and obtain the channel attention map. Then,this paper proposed a multiscale spatial attention module by combining multiscale feature extraction with spatial attention. Through this module,various positions of the channel-weighted feature map were weighted to obtain the spatial-weighted feature map. Finally,the feature map whose channels and spatial positions were weighted was input into the subsequent network for feature extraction and classification. Experimental results show that this method improves the expression recognition accuracy by 03% and 1%8% on the extended Cohn-Kanada and OULUCASIA NIR(near infrared)&VIS(visible light)datasets,respectively,which proves the effectiveness of this method.  
摘要:  
人脸表情识别一直是计算机视觉领域的一个难题.近年来,随着深度学习的飞速发展,一些基于卷积神经网络的方法大大提高了人脸表情识别的准确率,但未能充分利用人脸图像中的信息,这是由于对于面部表情识别有意义的特征主要集中在一些关键位置,例如眼睛、鼻子和嘴巴等区域,因此在特征提取时增加这些关键位置的权重可以改善表情识别的效果.为此,提出一种基于注意力机制的人脸表情识别网络.首先在主干网络中加入了深浅层特征融合结构,以充分提取原始图像中不同尺度的浅层特征,并将其与深层特征级联,以减少前向传播时的信息丢失.然后在网络中嵌入一种基于两步法的通道注意力模块,对级联后的特征图中的通道信息进行编码,得到通道注意力图,再将其与级联特征图逐元素相乘,得到通道加权特征图,将多尺度特征提取与空间注意力相结合,提出多尺度空间注意力模块,对通道加权特征图的不同位置进行加权,得到空间加权特征图.最后将通道和空间均已加权的特征图输入到后续网络中继续进行特征提取和分类.实验结果表明,所提出的方法与现有的基于深度学习的方法相比,在扩展的Cohn-Kanada数据集上的表情识别准确率提高了0~3%,在OULU-CASIA NIR&VIS数据集上的表情识别准确率提高了1%~8%,证明了该方法的有效性.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0493-2137  
入藏号:  
CSCD:7200280  
Indexed Date:  
2022-07-08  
Institution Address:  
天津大学微电子学院, 天津 300072, 中国  
Institution Address:  
School of Microelectronics,Tianjin University, Tianjin 300072, China  
Institution:  
天津大学微电子学院  
Institution:  
School of Microelectronics,Tianjin University

### 129. MFIHNet: Multiscale Feature Interaction Hybrid Network for Change

摘要: Remote sensing image change detection (RSCD) based on deep learning technology has made remarkable achievements. Meanwhile, the enhancement of network architectures and advancements in optimization algorithms have pushed RSCD performance to a higher stage. However, the existing RSCD methods mainly focus on extracting differential information between pixel pairs from bitemporal images, while neglecting the importance of using complementary multiscale features to uncover hierarchical semantic change information. To address these issues, we propose a multiscale feature interaction hybrid network (MFIHNet), which aims to enhance feature discriminability by multiscale features interaction and fusion to enhance model RSCD performance. Specifically, we first design a cascaded convolutional neural networks (CNN)-Transformer feature extraction network to capture hierarchical features at different scales. This strategy enables the network to preserve detailed information in shallow layers while grasping more contextual information in high layers. Subsequently, based on the differences between hierarchical features, we design a novel edge enhancement module (EM) to adaptively focus on key areas under the guidance of edge information to make the changed information clearer. Furthermore, to ensure complementary advantages among different feature layers, we devise a novel cross-scale feature interaction module, which introduces a region-specific atrous convolution into the multiscale attention mechanism for improving feature coassistance capacity. In this way, the MFIHNet not only effectively obtains different types of fine-grained information but also reduces the loss incurred during the feature fusion process, thereby improving the performance of remote sensing RSCD tasks. Extensive experimental results on the challenging CDD and GZ-CD datasets, with mean F1 scores reaching 98.1% and 87.4%, respectively, demonstrate that the proposed method achieves competitive performance.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001322151100015  
Indexed Date:  
2024-10-10  
Institution Address:  
Beijing Informat Sci & Technol Univ, Beijing 100029, Peoples R China  
Institution Address:  
Beijing Remote Sensing Informat Inst, Beijing 100083, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100045, Peoples R China  
Institution Address:  
Northeastern Univ, Shenyang 110819, Liaoning, Peoples R China  
Institution Address:  
Harbin Inst Technol, Harbin 150001, Peoples R China  
Institution:  
Institution:  
Beijing Remote Sensing Informat Inst  
Institution:  
Institution:  
Institution:

### 130. Improved YOLOv5 Object Detection Algorithm for Remote Sensing Images

摘要: An improved YOLOv5 is proposed to address complex backgrounds and small objects missing detection in remote sensing images. Firstly, considering that the high-level feature map contains little small object information caused by down-sampling of convolutional neural networks, low-level feature is reused to increase the small target feature information. The EMFFN(efficient multi-scale feature fusion network) is used in the feature fusion stage instead of the original PANet(path aggregation network) to efficiently fuse the feature map information at different scales by adding jump connections and skip connections. Finally, a bidirectional feature attention mechanism(BFAM) including channels attention and pixel attention is designed to improve detection in complex background. To evaluate the proposed model, this paper uses two remote sensing image datasets, DIOR and RSOD. The experimental results show that the improved YOLOv5 model achieves 87.8% and 96.6% detection accuracy in the DIOR and RSOD datasets respectively, which is 5.2 and 1.6 percentage points better than the original YOLOv5 algorithm, effectively improving the detection accuracy of small targets in complex backgrounds.  
摘要:  
针对遥感影像目标检测中复杂背景的干扰,小目标检测效果差等问题,提出一种改进YOLOv5(you only look once v5)的遥感影像目标检测模型。针对卷积神经网络下采样导致的特征图中包含的小目标信息较少或消失的问题,引入特征复用以增加特征图中的小目标特征信息;在特征融合阶段时使用EMFFN(efficient multi-scale feature fusion network)的特征融合网络代替原有的PANet(path aggregation network),通过添加跳跃连接以及跨层连接高效融合不同尺度的特征图信息;为了应对复杂背景带来的检测效果变差的问题,提出了一种包含通道与像素的双向特征注意力机制(bidirectional feature attention mechanism,BFAM),以提高模型在复杂背景下的检测效果。实验结果表明,改进后的YOLOv5模型在DIOR数据集与RSOD数据集中分别取得了87.8%和96.6%的检测精度,相较原算法分别提高5.2和1.6个百分点,有效提高了复杂背景下的小目标检测精度。  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
4  
被引频次合计:  
4  
ISSN:  
1002-8331  
入藏号:  
CSCD:7544074  
Indexed Date:  
2023-12-08  
Institution Address:  
宁夏大学地理科学与规划学院, 银川, 宁夏 750021, 中国  
Institution Address:  
School of Geography and Planning,Ningxia University, Yinchuan, Ningxia 750021, China  
Institution:  
宁夏大学地理科学与规划学院  
Institution:  
School of Geography and Planning,Ningxia University

### 131. Multiscale Spatial-Spectral Feature Extraction Network for Hyperspectral

摘要: Convolutional neural networks have garnered increasing interest for the supervised classification of hyperspectral imagery. However, images with a wide variety ofspatial land-cover sizes can hinder the feature-extraction ability of traditional convolutional networks. Consequently, many approaches intended to extract multiscale features have emerged; these techniques typically extract features in multiple parallel branches using convolutions of differing kernel sizes with concatenation or addition employed to fuse the features resulting from the various branches. In contrast, the present work explores a multiscale spatial-spectral feature-extraction network that operates in a more granular manner. Specifically, in the proposed network, a multibranch structure expands the convolutional receptive fields through the partitioning of input feature maps, applying hierarchical connections across the partitions, crosschannel feature fusion via pointwise convolution, and depthwise three-dimensional (3-D) convolutions for feature extraction. Experimental results reveal that the proposed multiscale spatial-spectral feature-fusion network outperforms other state-of-the-art networks at the supervised classification of hyperspectral imagery while being robust to limited training data.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000814673000002  
Indexed Date:  
2022-06-29  
Institution Address:  
Changan Univ, Sch Elect & Control Engn, Xian 710064, Peoples R China  
Institution Address:  
Mississippi State Univ, Dept Elect & Comp Engn, Mississippi State, MS 39762 USA  
Institution:  
Institution:

### 132. MSFFA: a multi-scale feature fusion and attention mechanism network for

摘要: Crowd counting has been a growing hot topic in the computer vision community in recent years due to its extensive applications in the fields of public safety and commercial planning. However, up to now, it has been still a challenging task in realistic scenes owing to large-scale variations and complex background interference. In this paper, we have proposed an efficient end-to-end Multi-Scale Feature Fusion and Attention mechanism CNN network, named as MSFFA. The presented network consists of three parts: the front-end of the low-level feature extractor, the mid-end of the multi-scale feature fusion operator and the back-end of the density map generator. Among them, most significantly, in the mid-end, we stack three MSFF blocks with the residual connection, which on the one hand, makes the network obtain large-scale continuous variations and on the other hand, enhances the information transmission. Meanwhile, a global attention mechanism module is employed to extract effective features in complex background scenes. Our method has been evaluated on three public datasets, including ShanghaiTech, UCF-QNRF and UCF\_CC\_50. Experimental results show that our method outperforms some existing advanced approaches, indicating its excellent accuracy and stability.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
10  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
10  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:000741931800001  
Indexed Date:  
2022-01-20  
Institution Address:  
Peoples Publ Secur Univ China, Coll Informat & Cyber Secur, Beijing 102600, Peoples R China  
Institution:

### 133. HAMNet: hyperspectral image classification based on hybrid neural

摘要: Recently, convolutional neural network (CNN) has made great progress in hyperspectral image (HSI) classification. Considering the problems of high dimensions, limited training samples and intra-class variations of hyperspectral data, there are challenges for traditional pure 2D or 3D deep convolutional neural networks in classifying HSI. Deeper layers bring gradient dispersion, while 3D feature blocks bring a large number of parameters during feature fusion. In this paper, an end-to-end hybrid convolutional neural network is proposed for HSI classification. Firstly, 3D, 2D and 1D convolution modules are applied, respectively, to perform joint feature extraction of spatial and spectral information. Secondly, a new 3D multi-scale feature fusion strategy is proposed to fuse the high-level and low-level features for ensuring the feature sufficiency. Moreover, channel attention mechanism is introduced to avoid feature channel redundancy and strengthen effective features. Comparative experimental results show that the method can receive satisfactory results on public data sets and small-sample learning problem.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:000841661900001  
Indexed Date:  
2022-08-25  
Institution Address:  
Qingdao Univ Sci & Technol, Coll Electromech Engn, Qingdao 266061, Peoples R China  
Institution Address:  
Northwest A&F Univ, Coll Mech & Elect Engn, Yangling, Shaanxi, Peoples R China  
Institution Address:  
Qingdao Univ Sci & Technol, Sch Math & Phys, Qingdao, Peoples R China  
Institution:  
Institution:  
Institution:

### 134. Image Manipulation Localization Using Multi-Scale Feature Fusion and

摘要: Image manipulation localization is a technique that can efficiently segment the tampered regions from a suspicious image. Existing work usually trains a detection model by fusing the features from diverse data streams, e.g., noise inconsistency, recompression inconsistency, and local inconsistency. They, however, ignore a fact that not all tampered images contain these data streams. As a result, high feature redundancy may cause a large number of false detection for tampered region. To address this problem, this paper designs an end-to-end high-confidence localization network architecture. First, deep convolutional neural networks are utilized to extract multi-scale feature sets from the RGB streams. We then design a semantic refined bi-directional feature integration module to fully fuse multi-scale adjacent features and significantly enhance feature representation. Subsequently, morphological operations are introduced to extract multi-scale edge information, which can efficiently reduce feature redundancy by generating wider high-resolution edges during image reconstructing. Finally, a deep semantic residual decoder is sequentially re-constructed by spreading deep semantic information into each decoding stage. The proposed method can not only improve the manipulation localization accuracy, but also guarantee the model robustness. Extensive experiments demonstrate that our method can obtain an effective performance in locating forged regions over different large-scale image sets, and outperforms most of state-of-the-art methods with higher localization accuracy and stronger robustness.  
Web of Science 核心合集中的 "被引频次":  
14  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
14  
ISSN:  
1520-9210  
eISSN:  
1941-0077  
入藏号:  
WOS:001121212400019  
Indexed Date:  
2024-02-05  
Institution Address:  
Shanghai Univ Elect Power, Coll Comp Sci & Technol, Shanghai 201306, Peoples R China  
Institution Address:  
Guangxi Normal Univ, Guangxi Key Lab Multisource Informat Min & Secur, Guilin 541004, Peoples R China  
Institution Address:  
Fudan Univ, Sch Comp Sci, Shanghai 200433, Peoples R China  
Institution Address:  
Univ Shanghai Sci & Technol, Sch Opt & Comp Engn, Shanghai 200093, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 135. Multilevel Attention and Multiscale Feature Fusion Network for Author

摘要: How to effectively extract features with high representation ability has always been a research topic and a challenge for classification tasks. Most of the existing methods mainly solve the problem by using deep convolutional neural networks as feature extractors. Although a series of excellent network structures have been successful in the field of Chinese ink-wash painting classification, but most of them adopted the methods of only simple augmentation of the network structures and direct fusion of different scale features, which limit the network to further extract semantically rich and scale-invariant feature information, thus hindering the improvement of classification performance. In this paper, a novel model based on multi-level attention and multi-scale feature fusion is proposed. The model extracts three types of feature maps from the low-level, middle-level and high-level layers of the pretrained deep neural network firstly. Then, the low-level and middle-level feature maps are processed by the spatial attention module, nevertheless the high-level feature maps are processed by the scale invariance module to increase the scale-invariance properties. Moreover, the conditional random field module is adopted to fuse the optimized three-scale feature maps, and the channel attention module is followed to refine the features. Finally, the multi-level deep supervision strategy is utilized to optimize the model for better performance. To verify the effectiveness of the model, extensive experimental results on the Chinese ink-wash painting dataset created in this work show that the classification performance of the model is better than other mainstream research methods.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
1026-0226  
eISSN:  
1607-887X  
入藏号:  
WOS:000773703800002  
Indexed Date:  
2022-04-17  
Institution Address:  
Changzhou Inst Technol, Sch Comp Informat & Engn, Changzhou, Peoples R China  
Institution:

### 136. Graph Convolution Based Cross-Network Multiscale Feature Fusion for Deep

摘要: Vessel segmentation is widely used to help with vascular disease diagnosis. Vessels reconstructed using existing methods are often not sufficiently accurate to meet clinical use standards. This is because 3D vessel structures are highly complicated and exhibit unique characteristics, including sparsity and anisotropy. In this paper, we propose a novel hybrid deep neural network for vessel segmentation. Our network consists of two cascaded subnetworks performing initial and refined segmentation respectively. The second subnetwork further has two tightly coupled components, a traditional CNN-based U-Net and a graph U-Net. Cross-network multi-scale feature fusion is performed between these two U-shaped networks to effectively support high-quality vessel segmentation. The entire cascaded network can be trained from end to end. The graph in the second subnetwork is constructed according to a vessel probability map as well as appearance and semantic similarities in the original CT volume. To tackle the challenges caused by the sparsity and anisotropy of vessels, a higher percentage of graph nodes are distributed in areas that potentially contain vessels while a higher percentage of edges follow the orientation of potential nearby vessels. Extensive experiments demonstrate our deep network achieves state-of-the-art 3D vessel segmentation performance on multiple public and in-house datasets.  
在 BIOSIS Citation Index 中的被引频次:  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
24  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
24  
ISSN:  
0278-0062  
eISSN:  
1558-254X  
入藏号:  
WOS:000907160700015  
PubMed ID:  
36112564  
Indexed Date:  
2023-01-25  
Institution Address:  
Univ Hong Kong, Dept Comp Sci, Hong Kong, Peoples R China  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Artificial Intelligence, Pattern Recognit & Intelligent Syst Lab, Beijing 100876, Peoples R China  
Institution Address:  
Beihang Univ, Inst Artificial Intelligence, Beijing 100191, Peoples R China  
Institution Address:  
AI Lab, Deepwise Healthcare, Beijing 100081, Peoples R China  
Institution Address:  
Zhejiang Univ, Sch Med, Dept Cardiol, Hangzhou 310009, Peoples R China  
Institution Address:  
Zhejiang Univ, Affiliated Hosp 2, Sch Med, Cardiovasc Key Lab Zhejiang Prov, Hangzhou 310009, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
AI Lab  
Institution:  
Institution:

### 137. PVT-SAR: An Arbitrarily Oriented SAR Ship Detector With Pyramid Vision

摘要: The development of deep learning has significantly boosted the development of ship detection in synthetic aperture radar (SAR) images. Most previous works rely on the convolutional neural networks (CNNs), which extract characteristics through local receptive fields and are sensitive to noise. Moreover, these detectors have limited performance in large-scale and complex scenes due to the strong interference of inshore background and the variability of target imaging characteristics. In this article, a novel SAR ship detection framework is proposed, which establishes the pyramid vision transformer (PVT) paradigm for multiscale feature representations in SAR images and, hence, is referred to as PVT-SAR. It breaks the limitation of the CNN receptive field and captures the global dependence through the self-attention mechanism. Since the difficulties of object detection in SAR and natural images are quite different, directly applying the existing transformer structure, such as PVT-small, cannot achieve satisfactory performance for SAR object detection. Compared with the PVT, overlapping patch embedding and mixed transformer encoder modules are incorporated to overcome the problems of densely arranged targets and insufficient data. Then, a multiscale feature fusion module is designed to further improve the detection ability for small targets. Moreover, a normalized Gaussian Wasserstein distance loss is employed to suppress the influence of scattering interference at the ship's boundary. The superiority of the proposed PVT-SAR detector over several state-of-the-art-oriented bounding box detectors has been evaluated in both inshore and offshore scenes on two commonly used SAR ship datasets (i.e., RSSDD and HRSID).  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
23  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
24  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000894925400023  
Indexed Date:  
2023-01-05  
Institution Address:  
Shanghai Jiao Tong Univ, Sch Elect Informat & Elect Engn, Shanghai 200240, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Dept Comp Sci & Engn, Shanghai, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, AI Inst, MoE Key Lab Artificial Intelligence, Shanghai 200240, Peoples R China  
Institution Address:  
Beijing Inst Remote Sensing Informat, Beijing 100192, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Beijing Inst Remote Sensing Informat

### 138. Automatic feature extraction and fusion recognition of motor imagery EEG

摘要: A motor imagery EEG (MI-EEG) signal is often selected as the driving signal in an active brain computer interface (BCI) system, and it has been a popular field to recognize MI-EEG images via convolutional neural network (CNN), which poses a potential problem for maintaining the integrity of the time-frequency-space information in MI-EEG images and exploring the feature fusion mechanism in the CNN. However, information is excessively compressed in the present MI-EEG image, and the sequential CNN is unfavorable for the comprehensive utilization of local features. In this paper, a multidimensional MI-EEG imaging method is proposed, which is based on time-frequency analysis and the Clough-Tocher (CT) interpolation algorithm. The time-frequency matrix of each electrode is generated via continuous wavelet transform (WT), and the relevant section of frequency is extracted and divided into nine submatrices, the longitudinal sums and lengths of which are calculated along the directions of frequency and time successively to produce a 3 x 3 feature matrix for each electrode. Then, feature matrix of each electrode is interpolated to coincide with their corresponding coordinates, thereby yielding a WT-based multidimensional image, called WTMI. Meanwhile, a multilevel and multiscale feature fusion convolutional neural network (MLMSFFCNN) is designed for WTMI, which has dense information, low signal-to-noise ratio, and strong spatial distribution. Extensive experiments are conducted on the BCI Competition IV 2a and 2b datasets, and accuracies of 92.95% and 97.03% are yielded based on 10-fold cross-validation, respectively, which exceed those of the state-of-the-art imaging methods. The kappa values and p values demonstrate that our method has lower class skew and error costs. The experimental results demonstrate that WTMI can fully represent the time-frequency-space features of MI-EEG and that MLMSFFCNN is beneficial for improving the collection of multiscale features and the fusion recognition of general and abstract features for WTMI.  
Web of Science 核心合集中的 "被引频次":  
27  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
7  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
30  
ISSN:  
0140-0118  
eISSN:  
1741-0444  
入藏号:  
WOS:000687513900001  
PubMed ID:  
34424453  
Indexed Date:  
2021-08-28  
Institution Address:  
Beijing Univ Technol, Fac Informat Technol, Beijing 100124, Peoples R China  
Institution Address:  
Beijing Key Lab Computat Intelligence & Intellige, Beijing 100124, Peoples R China  
Institution Address:  
Minist Educ, Engn Res Ctr Digital Community, Beijing 100124, Peoples R China  
Institution:  
Institution:  
Beijing Key Lab Computat Intelligence & Intellige  
Institution:  
Minist Educ

### 139. An Efficient Center-Based Method With Multilevel Auxiliary Supervision

摘要: The problem of multiscale ship detection in synthetic aperture radar (SAR) images has received much attention with the development of deep convolutional neural networks (DCNNs). However, existing DCNN-based multiscale SAR ship detection methods often lead to time-consuming detection process due to the massive parameters therein. To address this issue, a lightweight center-based detector with the multilevel auxiliary supervision (MLAS) structure is proposed in this article. First, an extremely lightweight backbone network is designed to improve the computation efficiency and extract SAR image features in a bottom-up manner. Then, a feature fusion network containing three multiscale feature fusion modules is introduced to combine semantic features with different levels. Finally, a novel MLAS-based framework is proposed to train our DCNN with multilevel auxiliary detection subnets. MLAS improves the performance of multiscale ship detection benefiting from the guidance of multilevel attention. Experimental results on the open SAR image dataset SSDD show that our proposed detector achieves a similar average precision for the problem of multiscale SAR ship detection but significantly reduces the computation burden of state-of-the-art methods. The required number of floating points of operations of our method is only 21.70%, 19.30%, and 4.81% of those of CenterNet, YOLOv3, and RetinaNet, respectively, and the number of learnable weights in our method is only 0.68 million that is 5.63%, 1.10%, 2.98% of those of the aforementioned three existing methods, respectively.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000849263100009  
Indexed Date:  
2022-09-10  
Institution Address:  
Tsinghua Univ, Dept Elect Engn, Beijing 100084, Peoples R China  
Institution Address:  
Tsinghua Univ, Shenzhen Int Grad Sch, Shenzhen 518055, Peoples R China  
Institution:  
Institution:

### 140. Traffic sign detection based on multi-scale feature extraction and

摘要: Existing algorithms have difficulty in solving the two tasks of localization and classification simultaneously when performing traffic sign detection on realistic images of complex traffic scenes. In order to solve the above problems, a new road traffic sign dataset is created, and based on the YOLOv4 algorithm, for the complexity of realistic traffic scene images and the large variation in the size of traffic signs in the images, the multi-scale feature extraction module, cascade feature fusion module and attention mechanism module are designed to improve the algorithm's ability to locate and classify traffic signs simultaneously. Experimental results on the newly created dataset show that the improved algorithm achieves a mean average precision of 84.44%, which is higher than several major CNN-based object detection algorithms for the same type of task.  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
16  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
17  
ISSN:  
0920-8542  
eISSN:  
1573-0484  
入藏号:  
WOS:000836828400002  
Indexed Date:  
2022-08-15  
Institution Address:  
Hefei Univ Technol, Sch Comp Sci & Informat Engn, Hefei 230009, Peoples R China  
Institution Address:  
Anhui Mine IOT & Secur Monitoring Technol Key Lab, Hefei 230088, Peoples R China  
Institution Address:  
Hefei Univ Technol, Engn Res Ctr Safety Crit Ind Measurement & Contro, Minist Educ, Hefei 230009, Peoples R China  
Institution Address:  
HeFei Univ Technol, Intelligent Mfg Inst, Hefei 230009, Peoples R China  
Institution:  
Institution:  
Anhui Mine IOT & Secur Monitoring Technol Key Lab  
Institution:  
Institution:

### 141. Ultralightweight Spatial-Spectral Feature Cooperation Network for Change

摘要: Deep convolutional neural networks (CNNs) have achieved much success in remote sensing image change detection (CD) but still suffer from two main problems. First, the existing multiscale feature fusion methods often use redundant feature extraction and fusion strategies, which often lead to high computational costs and memory usage. Second, the regular attention mechanism in CD is difficult to model spatial-spectral features and generate 3-D attention weights at the same time, ignoring the cooperation between spatial features and spectral features. To address the above issues, an efficient ultralightweight spatial-spectral feature cooperation network (USSFC-Net) is proposed for CD in this article. The proposed USSFC-Net has two main advantages. First, a multiscale decoupled convolution (MSDConv) is designed, which is clearly different from the popular atrous spatial pyramid pooling (ASPP) module and its variants since it can flexibly capture the multiscale features of changed objects using cyclic multiscale convolution. Meanwhile, the design of MSDConv can greatly reduce the number of parameters and computational redundancy. Second, an efficient spatial-spectral feature cooperation (SSFC) strategy is introduced to obtain richer features. The SSFC differs from the existing 2-D attention mechanisms since it learns 3-D spatial-spectral attention weights without adding any parameters. The experiments on three datasets for remote sensing image CD demonstrate that the proposed USSFC-Net achieves better CD accuracy than most CNNs-based methods and requires lower computational costs and fewer parameters, even it is superior to some Transformer-based methods. The code is available at https://github.com/SUST-reynole/USSFC-Net.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
63  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
65  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000965182200009  
Indexed Date:  
2023-05-06  
Institution Address:  
Shaanxi Univ Sci & Technol, Shaanxi Joint Lab Artificial Intelligence, Xian 710021, Peoples R China  
Institution Address:  
Univ Posts & Telecommun, Sch Comp Sci & Technol, Xian 710121, Peoples R China  
Institution Address:  
Univ Technol, Sch Comp Sci & Engn, Xian 710048, Peoples R China  
Institution Address:  
Xidian Univ, Int Res Ctr Intelligent Percept & Computat, Key Lab Intelligent Percept & Image Understanding, Minist Educ, Xian 710071, Peoples R China  
Institution Address:  
Bielefeld Univ, Fac Technol, D-33619 Bielefeld, Germany  
Institution Address:  
Brunel Univ London, Dept Elect & Elect Engn, Uxbridge UB8 3PH, England  
Institution Address:  
Xi An Jiao Tong Univ, Sch Mech Engn, Xian 710049, Peoples R China  
Institution:  
Institution:  
Univ Posts & Telecommun  
Institution:  
Univ Technol  
Institution:  
Institution:  
Institution:  
Institution:

### 142. Flotation Dosing State Recognition Based on Multiscale CNN Features and

摘要: To address the problems associated with online detection, low recognition efficiency, and strong subjectivity of the flotation dosing state, this paper proposes a flotation dosing state recognition method based on multiscale convolutional neural network (CNN) features and ranks automatic encoder kernel extreme learning machine (RAE-KELM). First, the flotation foam image is subjected to non-subsampled Shearlet multiscale decomposition, and the CNN is used to extract the depth features of each scale image and perform multiscale feature fusion. Then, the RAE-KELM is constructed, and an improved bacterial foraging algorithm based on quantum computing is used to optimize the RAE-KELM parameters. Finally, the optimal RAE-KELM model is obtained through self-built dataset training to realize the adaptive recognition of the flotation dosing state. The experimental results demonstrate that the recognition accuracy of the method can reach 98.88%. Additionally, the method reduces manual interventions, which can improve production efficiency.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1006-4125  
入藏号:  
WOS:000686490800043  
Indexed Date:  
2021-08-29  
Institution Address:  
Fuzhou Univ, Coll Phys & Informat Engn, Fuzhou 350108, Fujian, Peoples R China  
Institution:

### 143. Multiscale and Cross-Level Attention Learning for Hyperspectral Image

摘要: Transformer-based networks, which can well model the global characteristics of inputted data using the attention mechanism, have been widely applied to hyperspectral image (HSI) classification and achieved promising results. However, the existing networks fail to explore complex local land cover structures in different scales of shapes in hyperspectral remote sensing images. Therefore, a novel network named multiscale and cross-level attention learning (MCAL) network is proposed to fully explore both the global and local multiscale features of pixels for classification. To encounter local spatial context of pixels in the transformer, a multiscale feature extraction (MSFE) module is constructed and implemented into the transformer-based networks. Moreover, a cross-level feature fusion (CLFF) module is proposed to adaptively fuse features from the hierarchical structure of MSFEs using the attention mechanism. Finally, the spectral attention module (SAM) is implemented prior to the hierarchical structure of MSFEs, by which both the spatial context and spectral information are jointly emphasized for hyperspectral classification. Experiments over several benchmark datasets demonstrate that the proposed MCAL obviously outperforms both the convolutional neural network (CNN)-based and transformer-based state-of-the-art networks for hyperspectral classification.  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
61  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
62  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000992306700004  
Indexed Date:  
2023-06-10  
Institution Address:  
Northwestern Polytech Univ, Sch Elect & Informat, Xian 710129, Peoples R China  
Institution Address:  
Shanghai Inst Satellites Engn, Key Lab Millimeter Wave Imaging Technol, Shanghai 201100, Peoples R China  
Institution:  
Institution:  
Shanghai Inst Satellites Engn

### 144. BAF-Detector: An Efficient CNN-Based Detector for Photovoltaic Cell

摘要: The multiscale defect detection for photovoltaic (PV) cell electroluminescence (EL) images is a challenging task, due to the feature vanishing as network deepens. To address this problem, an attention-based top-down and bottom-up architecture is developed to accomplish multiscale feature fusion. This architecture, called bidirectional attention feature pyramid network (BAFPN), can make all layers of the pyramid share similar semantic features. In BAFPN, cosine similarity is employed to measure the importance of each pixel in the fused features. Furthermore, a novel object detector is proposed, called BAF-Detector, which embeds BAFPN into region proposal network in Faster RCNN+FPN. BAFPN improves the robustness of the network to scales, thus the proposed detector achieves a good performance in multiscale defects detection task. Finally, the experimental results on a large-scale EL dataset, including 3629 images, 2129 of which are defective, show that the proposed method achieves 98.70% (F-measure), 88.07% (mAP), and 73.29% (IoU) in terms of multiscale defects classification and detection results in raw PV cell EL images.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
8  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
92  
被引频次合计:  
101  
ISSN:  
0278-0046  
eISSN:  
1557-9948  
入藏号:  
WOS:000728182300097  
Indexed Date:  
2022-09-03  
Institution Address:  
Beihang Univ, Sch Comp Sci & Engn, State Key Lab Virtual Real Technol & Syst, Beijing 100191, Peoples R China  
Institution Address:  
Hebei Univ Technol, Sch Artificial Intelligence & Data Sci, Tianjin 300130, Peoples R China  
Institution:  
Institution:

### 145. Research on YOLOv5s Improved Algorithm for Pavement Crack Detection in

摘要: Aiming to solve the problem of pavement crack detection in complex road environments, an improved algorithm based on YOLOv5s is proposed. First, the CBAM (Convolutional Block Attention Module) is introduced after the backbone network's C3 modules to focus the complex scene's practical information and enhance the model's attention to the crack region. Second, the BiFPN (Bidirectional Feature Pyramid Network) is used in Neck to replace the bidirectional PANet (Path Aggregation Network) in YOLOv5s to improve the multiscale feature fusion, which reduces detection leakage due to illumination and scale factors. Thirdly, the P6 detection head for denser cracks is added to the Head, and the CA (Coordinate Attention) module is introduced to improve the crack detection capability at multiple scales. Finally, the algorithm is experimentally compared with SSD (Single Shot MultiBox Detector), Faster-RCNN (Faster-Region Convolutional Neural Network), and unimproved YOLOv5s on the constructed pavement crack image dataset. The results show that compared with other algorithms, when mAP@0.5 (mean Average Precision when IoU=0.5) values are improved by 10.6%, 9.2%, and 4.6%, respectively, and cracks are identified better than the other three models.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001311189300001  
Indexed Date:  
2024-09-27  
Institution Address:  
Beibu Gulf Univ, Adv Sci & Technol Res Inst, Qinzhou 535011, Peoples R China  
Institution:

### 146. Change Guiding Network: Incorporating Change Prior to Guide Change

摘要: The rapid advancement of automated artificial intelligence algorithms and remote sensing instruments has benefited change detection (CD) tasks. However, there is still a lot of space to study for precise detection, especially the edge integrity and internal holes phenomenon of change features. In order to solve these problems, we design the change guiding network (CGNet) to tackle the insufficient expression problem of change features in the conventional U-Net structure adopted in previous methods, which causes inaccurate edge detection and internal holes. Change maps from deep features with rich semantic information are generated and used as prior information to guide multiscale feature fusion, which can improve the expression ability of change features. Meanwhile, we propose a self-attention module named change guide module, which can effectively capture the long-distance dependency among pixels and effectively overcomes the problem of the insufficient receptive field of traditional convolutional neural networks. On four major CD datasets, we verify the usefulness and efficiency of the CGNet, and a large number of experiments and ablation studies demonstrate the effectiveness of CGNet.  
Web of Science 核心合集中的 "被引频次":  
40  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
40  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001070732700013  
Indexed Date:  
2023-10-12  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & Re, Wuhan 430079, Peoples R China  
Institution Address:  
Univ Tokyo, Grad Sch Frontier Sci, Chiba 2778561, Japan  
Institution:  
Institution:

### 147. Tool wear prediction using multi-sensor data fusion and attention-based

摘要: Tool wear is significant in the machining process due to its direct relationship with part accuracy and quality. A precise prediction of tool wear can reduce downtime and improve product quality. Conventional methods of measuring tool wear are not feasible in the age of Industry 4.0. Therefore, this paper introduces a new attention-based deep learning model that accurately predicts tool wear. The model achieves multiscale feature fusion, focusing on the most relevant features. We proposed a channel Attention mechanism that integrates with residual connections, considering the weight of each feature map to improve the model performance. We conducted lathe-turning experiments under varying cutting conditions and collected multi-sensor data to assess the performance of the model and compare it with existing models such as Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), and Support Vector Regression (SVR). Model performance was evaluated using root mean square error (RMSE), mean absolute error (MAE), and the coefficient of determination (R2). The results demonstrated that the proposed model outperformed the others across all evaluation metrics.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0944-6524  
eISSN:  
1863-7353  
入藏号:  
WOS:001361430500001  
Indexed Date:  
2024-12-03  
Institution Address:  
Natl Inst Technol Warangal, Hanamkonda, Telangana, India  
Institution:

### 148. MSSA-Net: A novel multi-scale feature fusion and global self-attention

摘要: In medical image segmentation tasks, it is typical to adopt convolutional neural networks with a serial encoder-decoder structure. However, mainstream networks cannot simultaneously achieve sufficient extraction of global features and the fusion of multi-scale information, which may lead to unpromising results for the segmentation of pathological images. Therefore, this article proposed a novel multi-scale feature fusion and global self-attention network (MSSA-Net) for medical image segmentation. Specifically, we designed a parallel double-encoder network with a multi-scale feature fusion encoder (MS-Encoder) and a self-attention encoder (SA-Encoder). The SA-Encoder introduces the transformer's global self-attention mechanism to extract global features, and the MS-Encoder adopts atrous spatial pyramid pooling (ASPP) to realize multi-scale fusion. We have evaluated the proposed MSSA-Net using three medical segmentation datasets, covering various imaging modalities such as colonoscopy and magnetic resonance imaging. Experiments on the CVC-ClinicDC, the 2015 MICCAI subchallenge on automatic polyp detection dataset, and anatomical tracings of lesions after stroke (ATLAS) show that our MSSA-Net outperforms mainstream methods such as DoubleU-Net and TransUNet. Moreover, MSSA-Net can predict more accurate segmentation masks, especially in the case of ATLAS, which has challenging images such as multiple shadow areas and discrete lesions.  
在 BIOSIS Citation Index 中的被引频次:  
3  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
ISSN:  
1532-0626  
eISSN:  
1532-0634  
入藏号:  
WOS:000798632700001  
Indexed Date:  
2022-05-30  
Institution Address:  
Jimei Univ, Comp Engn Coll, Xiamen, Peoples R China  
Institution Address:  
Xiamen Med Coll, Affiliated Hosp 2, Xiamen, Peoples R China  
Institution:  
Institution:

### 149. Multi-Scale Feature Aggregation Network for Semantic Segmentation of

摘要: Land cover semantic segmentation is an important technique in land. It is very practical in land resource protection planning, geographical classification, surveying and mapping analysis. Deep learning shows excellent performance in picture segmentation in recent years, but there are few semantic segmentation algorithms for land cover. When dealing with land cover segmentation tasks, traditional semantic segmentation networks often have disadvantages such as low segmentation precision and weak generalization due to the loss of image detail information and the limitation of weight distribution. In order to achieve high-precision land cover segmentation, this article develops a multi-scale feature aggregation network. Traditional convolutional neural network downsampling procedure has problems of detail information loss and resolution degradation; to fix these problems, a multi-scale feature extraction spatial pyramid module is made to assemble regional context data from different areas. In order to address the issue of incomplete information of traditional convolutional neural networks at multiple sizes, a multi-scale feature fusion module is developed to fuse attributes from various layers and several sizes to boost segmentation accuracy. Finally, a multi-scale convolutional attention module is presented to enhance the segmentation's attention to the target in order to address the issue that the classic convolutional neural network has low attention capacity to the building waters in land cover segmentation. Through the contrast experiment and generalization experiment, it can be clearly demonstrated that the segmentation algorithm proposed in this paper realizes the high precision segmentation of land cover.  
Web of Science 核心合集中的 "被引频次":  
8  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
eISSN:  
2072-4292  
入藏号:  
WOS:000896525000001  
Indexed Date:  
2022-12-25  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Collaborat Innovat Ctr Atmospher Environm & Equipm, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Forestry Univ, Coll Informat Sci & Technol, Nanjing 210000, Peoples R China  
Institution:  
Institution:

### 150. Fused Adaptive Receptive Field Mechanism and Dynamic Multiscale Dilated

摘要: Side-scan sonar (SSS) is a vital sensor for marine survey, which is widely used in military and civilian fields. The accurate segmentation of SSS images is critical in sonar image intelligent interpretation. Existing SSS image segmentation methods have several limitations, such as insufficient feature extraction, relatively worse segmentation results for tiny target categories, and serious interference by seabed reverberation noise and bright shadow region. To overcome these issues, we propose a novel encoder-decoder architecture SSS image segmentation method based on convolution neural network (CNN). First, we extract the multiscale feature information contained in target region using the dynamic multiscale dilated convolution (DMDC\_Conv). Second, to further obtain the global and detail feature information, we construct the adaptive receptive field mechanism block (ARFM\_Block). Third, we design a feature fusion attention mechanism block (FFAM\_Block) to fuse high-level and low-level feature information with different scales and suppress background information interference. Final, we construct a tree structure optimization module (TSOM) to solve the problem of pixel misclassification and obtain refine SSS image segmentation results. Extensive experiments are carried out on the constructed real scene SSS image dataset. The experimental results show that the proposed method achieves 93.24% and 90.82% of mean pixel accuracy (MPA) and mean intersection over union (MIoU), respectively, which outperforms the other state-of-the-art methods and has a substantial advantage in inference speed and calculation parameters.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
14  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000855545400001  
Indexed Date:  
2022-01-01  
Institution Address:  
Xijing Univ, Sch Informat Engn, Xian 710123, Peoples R China  
Institution Address:  
Univ Queensland, Sch Earth & Environm Sci, St Lucia, Qld 4072, Australia  
Institution Address:  
Tianjin Univ Sci & Technol, Coll Artificial Intelligence, Tianjin 300457, Peoples R China  
Institution Address:  
Air Force Engn Univ, Sch Informat & Nav, Xian 710082, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 151. Defect classification on limited labeled samples with multiscale feature

摘要: Defect inspection is an essential part of ensuring the quality of industrial products. Deep learning has achieved great success in defect inspection when a large number of labeled samples are available. However, it is infeasible to collect and label numerous samples in many manufacturing processes. Meanwhile, deep learning methods cannot conform to the high defect recognition accuracy of strict production requirements when the labeled samples are scarce but varied. This paper proposed a novel convolutional neural network architecture and a semi-supervised learning strategy using soft pseudo labels and a mutual correction classifier to improve the defect inspection accuracy when labeled samples are scarce. The effectiveness of the proposed method is verified on a famous industrial defect inspection benchmark dataset and a practical dataset containing images collected from actual injection molding production lines. The results indicate that the proposed method achieves an accuracy of 99.03% on the benchmark defect dataset, which is approximately 13.2% higher than other methods when the training dataset contains only 45 labeled images and 135 unlabeled samples per category. The best accuracy on the benchmark dataset obtained by the proposed method reaches 99.72%. Besides, an average accuracy of 99.25% is achieved with only 20 labeled samples and 180 unlabeled samples per category in the practical defect inspection task. Visualization methods prove that the performance improvement comes from the proposed multiscale architecture and the semi-supervised learning strategy. The proposed method can be used in practical defect inspection applications of industrial manufacturing, such as steel rolling, welding, and injection molding.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
9  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000710613300001  
Indexed Date:  
2021-10-24  
Institution Address:  
Huazhong Univ Sci & Technol, Sch Mat Sci & Engn, State Key Lab Mat Proc & Die & Mould Technol, Wuhan 430074, Peoples R China  
Institution:

### 152. EFPN: Effective medical image detection using feature pyramid fusion

摘要: Feature pyramid networks (FPNs) are widely used in the existing deep detection models to help them utilize multi-scale features. However, there exist two multi-scale feature fusion problems for the FPN-based deep detection models in medical image detection tasks: insufficient multi-scale feature fusion and the same importance for multi-scale features. Therefore, in this work, we propose a new enhanced backbone model, EFPNs, to overcome these problems and help the existing FPN-based detection models to achieve much better medical image detection performances. We first introduce an additional top-down pyramid to help the detection networks fuse deeper multi-scale information; then, a scale enhancement module is developed to use different sizes of kernels to generate more diverse multi-scale features. Finally, we propose a feature fusion attention module to estimate and assign different importance weights to features with different depths and scales. Extensive experiments are conducted on two public lesion detection datasets for different medical image modalities (X-ray and MRI). On the mAP and mR evaluation metrics, EFPN-based Faster R-CNNs improved 1.55% and 4.3% on the PenD (X-ray) dataset, and 2.74% and 3.1% on the BraTs (MRI) dataset, respectively. EFPN-based Faster R-CNNs achieve much better performances than the state-of-the-art baselines in medical image detection tasks. The proposed three improvements are all essential and effective for EFPNs to achieve superior performances; and besides Faster R-CNNs, EFPNs can be easily applied to other deep models to significantly enhance their performances in medical image detection tasks.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
4  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
12  
被引频次合计:  
12  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:001034226300001  
PubMed ID:  
37348265  
Indexed Date:  
2023-08-10  
Institution Address:  
Hebei Univ Technol, State Key Lab Reliabil & Intelligence Elect Equipm, Tianjin, Peoples R China  
Institution Address:  
Hainan Women & Childrens Med Ctr, Dept Radiol, Haikou, Peoples R China  
Institution Address:  
TU Wien, Inst Log & Computat, Vienna, Austria  
Institution Address:  
Univ Oxford, Dept Comp Sci, Oxford, England  
Institution:  
Institution:  
Hainan Women & Childrens Med Ctr  
Institution:  
Institution:

### 153. Lightweight train image fault detection model based on location

摘要: In addition to the imperative for accurate detection, the computational and storage costs of defect detection models are often considered in computing environments with limited resources, such as mobile devices and edge nodes. Consequently, a lightweight fault detection network model (Light-LIENet), based on location information enhancement, is proposed. It is capable of achieving real-time and high-precision detection while minimizing computational overhead and parameter complexity. Light-LIENet uses a convolutional neural network as its foundational architecture for extracting pertinent image features. Furthermore, it integrates a location feature network (LFNet) to accomplish multiscale feature fusion and enhance object location information. LFNet includes a location sensitive block, a feature enhancement block, and a lightweight convolution block. These blocks collectively contribute to enhancing location information, augmenting feature extraction, and minimizing computational amount. Experimental results show that, when compared to stateof-the-art lightweight detectors on the freight train image fault dataset, Light-LIENet exhibits impressive performance. With a detection speed of 62 frames per second, a mere 1.611 million model parameters, and a computational requirement of just 3.984 G, it unequivocally demonstrates its effectiveness and practicality.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001416297600001  
Indexed Date:  
2025-02-14  
Institution Address:  
Hunan Univ Technol, Sch Comp Sci, Zhuzhou 412007, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Sch Software Engn, Zhuhai 519082, Peoples R China  
Institution:  
Institution:

### 154. Remote sensing landslide target detection method based on improved

摘要: To timely detect landslide hazards to start emergency rescue, an improved Faster R-CNN algorithm is proposed for remote sensing image landslide detection. First, the gamma transform and Gaussian filtering methods of image enhancement are used to improve the quality of the images. Second, the effect of batchsize size on the model is eliminated using the group normalization method. Finally, multiscale feature fusion is performed by adding a feature pyramid network structure to optimize the extracted landslide small target features, and then the backbone network is set as deep residual shrinkage network 50 to make the model more focused on information useful for landslide detection. The experimental results show that the improved model improves the accuracy rate as well as the average precision by 8.8% and 8.4%, respectively, compared with the unimproved Faster R-CNN, and compared with the first-stage models, such as you only look once version 4 and single-shot detector, which verify the superiority of the model in our study and can detect landslide targets well. (c) 2022 Society of Photo-Optical Instrumentation Engineers (SPIE)  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
5  
eISSN:  
1931-3195  
入藏号:  
WOS:000917345800023  
Indexed Date:  
2023-02-17  
Institution Address:  
Fuzhou Univ, Sch Econ & Management, Fuzhou, Peoples R China  
Institution:

### 155. High-Resolution SAR Image Classification Using Multi-Scale Deep Feature

摘要: The classification of high-resolution (HR) synthetic aperture radar (SAR) images is of great importance for SAR scene interpretation and application. However, the presence of intricate spatial structural patterns and complex statistical nature makes SAR image classification a challenging task, especially in the case of limited labeled SAR data. This paper proposes a novel HR SAR image classification method, using a multi-scale deep feature fusion network and covariance pooling manifold network (MFFN-CPMN). MFFN-CPMN combines the advantages of local spatial features and global statistical properties and considers the multi-feature information fusion of SAR images in representation learning. First, we propose a Gabor-filtering-based multi-scale feature fusion network (MFFN) to capture the spatial pattern and get the discriminative features of SAR images. The MFFN belongs to a deep convolutional neural network (CNN). To make full use of a large amount of unlabeled data, the weights of each layer of MFFN are optimized by unsupervised denoising dual-sparse encoder. Moreover, the feature fusion strategy in MFFN can effectively exploit the complementary information between different levels and different scales. Second, we utilize a covariance pooling manifold network to extract further the global second-order statistics of SAR images over the fusional feature maps. Finally, the obtained covariance descriptor is more distinct for various land covers. Experimental results on four HR SAR images demonstrate the effectiveness of the proposed method and achieve promising results over other related algorithms.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
12  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
12  
eISSN:  
2072-4292  
入藏号:  
WOS:000611954300001  
Indexed Date:  
2021-02-10  
Institution Address:  
Xidian Univ, Sch Elect Engn, Remote Sensing Image Proc & Fus Grp, Xian 710071, Peoples R China  
Institution Address:  
Xidian Univ, Natl Lab Radar Signal Proc, Xian 710071, Peoples R China  
Institution:  
Institution:

### 156. Multi-Scale Hierarchical Feature Fusion for Infrared Small-Target

摘要: Detecting small targets in infrared images presents significant challenges due to their tiny size and complex backgrounds, making this task a hotspot for research. Traditional methods rely on assumption-based modeling and manual design, struggling to handle the variability of real-world scenarios. Although convolutional neural networks (CNNs) increase robustness to diverse scenes with a data-driven paradigm, many CNN-based methods are insufficient in capturing fine-grained details necessary for small targets and are less effective during multi-scale feature fusion. To overcome these challenges, we propose the novel Wide-scale Gated Fully Fusion Network (WGFFNet) in this article, which contributes to infrared small-target detection (IRSTD). WGFFNet uses a classic encoder-decoder structure, where the designed stepped fusion block (SFB) embedded in the feature extraction stage captures finer local context across multiple scales during encoding, and along the decoding path, the multi-level features are progressively integrated by a Fully Gated Interaction (FGI) Module to enhance feature representation. The inclusion of a boundary difference loss further optimizes the edge details of targets. We conducted comprehensive experiments on two public infrared small-target datasets: SIRST-V2 and IRSTD-1k. Quantitative and qualitative results demonstrate that our WGFFNet outperforms representative methods when considering various evaluation metrics together, achieving an improved detection performance and computational efficiency for detecting small targets in infrared images.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001419489300001  
Indexed Date:  
2025-02-18  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 100049, Peoples R China  
Institution Address:  
Chinese Acad Sci, Changchun Inst Opt Fine Mech & Phys, State Key Lab Appl Opt, Changchun 130033, Peoples R China  
Institution:  
Institution:  
Institution:

### 157. Gesture recognition algorithm based on multi-scale feature fusion in

摘要: With the rapid development of sensor technology and artificial intelligence, the video gesture recognition technology under the background of big data makes human-computer interaction more natural and flexible, bringing richer interactive experience to teaching, on-board control, electronic games, etc. In order to perform robust recognition under the conditions of illumination change, background clutter, rapid movement, partial occlusion, an algorithm based on multi-level feature fusion of two-stream convolutional neural network is proposed, which includes three main steps. Firstly, the Kinect sensor obtains RGB-D images to establish a gesture database. At the same time, data enhancement is performed on training and test sets. Then, a model of multi-level feature fusion of two-stream convolutional neural network is established and trained. Experiments result show that the proposed network model can robustly track and recognize gestures, and compared with the single-channel model, the average detection accuracy is improved by 1.08%, and mean average precision (mAP) is improved by 3.56%. The average recognition rate of gestures under occlusion and different light intensity was 93.98%. Finally, in the ASL dataset, LaRED dataset, and 1-miohand dataset, recognition accuracy shows satisfactory performances compared to the other method.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
26  
被引频次合计:  
26  
ISSN:  
1751-9659  
eISSN:  
1751-9667  
入藏号:  
WOS:000903045100001  
Indexed Date:  
2023-01-12  
Institution Address:  
Wuhan Univ Sci & Technol, Key Lab Met Equipment & Control Technol, Minist Educ, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Res Ctr Biol Manipulator & Intelligent Measurement, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Inst Precis Mfg, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Hubei Key Lab Mech Transmiss & Mfg Engn, Wuhan, Peoples R China  
Institution Address:  
Univ Portsmouth, Sch Comp, Portsmouth, England  
Institution Address:  
Wuhan Univ Sci & Technol, Res Ctr Biol Manipulator & Intelligent Measurement, Wuhan 430081, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 158. Fusion of Deep Learning Models for Multi-View Image Classification

摘要: Advances in Convolutional Neural Networks (CNN) have demonstrated state of the art performance in the tasks of image classification and object detection over the past decade. While significant progress has been made in development of more efficient networks, the computational and memory requirements still exceed practical limits in many applications. Additionally, the pose variability in such applications requires even larger training datasets for the network to generalize to all possible scenarios. The goal of this work is to develop an architecture for fusion of multiple views of a single target to provide robust classification with a lightweight backbone network used across all agents. Motivated by approaches to ensemble learning, we demonstrate that multiple weak learners with computationally efficient networks can combine to enhance classification accuracy. Three methods of fusion are considered: decision fusion, feature fusion, and multi-scale feature fusion. A novel network architecture is developed and implemented for each approach then trained and evaluated using synthetic data. For the feature fusion models, a custom training scheme is developed to minimize classification error while maintaining a common feature extraction backbone across agents. This conforms to a distributed classification use case where each agent has no prior knowledge of its position relative to target. Finally, we discuss the requirements for shared data of each approach in the context of applications with limited communication bandwidth.  
会议名称:  
Conference on Signal Processing, Sensor/Information Fusion, and Target Recognition XXXII  
会议日期:  
MAY 01-03, 2023  
会议地点:  
Orlando, FL  
会议赞助商:  
SPIE  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0277-786X  
eISSN:  
1996-756X  
ISBN:  
978-1-5106-6210-0; 978-1-5106-6211-7  
入藏号:  
WOS:001033524700012  
Indexed Date:  
2023-09-03  
Institution Address:  
US Army DEVCOM AC, Picatinny Arsenal, NJ 07806 USA  
Institution:  
US Army DEVCOM AC

### 159. Robust coverless image steganography based on DenseUNet with multi-scale

摘要: Coverless image steganography (CIS) have attracted considerable attention due to their ability to resist steganalysis detection completely. However, most of the existing CIS methods are weak in robustness to geometric attacks, and it is difficult to achieve a balance between geometric attacks and non-geometric attacks. So, a robust coverless image steganography method based on DenseUNet with multi-scale feature fusion attention mechanism is proposed in this paper. At the sender, an end-to-end hash sequence generation model is designed by combining the DenseUNet network with the multi-scale feature fusion attention mechanism to extract the multi-scale CNN features of the images, and as serve matching benchmarks. Secondly, a hybrid loss function is introduced into the network model for network training to generate hash sequences with robust features. Finally, the secret information is segmented into equal-length segments, and the image whose hash sequence matches the secret information segment is selected as a stego-images using the inverted index. At the receiver, the secret information was recovered from the stego-images using the constructed network model. Experimental results show that the proposed method has stronger performance in terms of robustness and security compared with existing CIS schemes, and achieves enhanced robustness against both the geometric attacks and non-geometric attacks at four different datasets.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001284778800002  
Indexed Date:  
2024-08-13  
Institution Address:  
Lanzhou Univ Technol, Sch Comp & Commun, Lanzhou 730050, Gansu, Peoples R China  
Institution:

### 160. Multi-Scale Feature Fusion Point Cloud Object Detection Based on

摘要: Existing point cloud object detection algorithms struggle to effectively capture spatial features across different scales, often resulting in inadequate responses to changes in object size and limited feature extraction capabilities, thereby affecting detection accuracy. To solve this problem, we present a point cloud object detection method based on multi-scale feature fusion of the original point cloud and projection, which aims to improve the multi-scale performance and completeness of feature extraction in point cloud object detection. First, we designed a 3D feature extraction module based on the 3D Swin Transformer. This module pre-processes the point cloud using a 3D Patch Partition approach and employs a self-attention mechanism within a 3D sliding window, along with a downsampling strategy, to effectively extract features at different scales. At the same time, we convert the 3D point cloud to a 2D image using projection technology and extract 2D features using the Swin Transformer. A 2D/3D feature fusion module is then built to integrate 2D and 3D features at the channel level through point-by-point addition and vector concatenation to improve feature completeness. Finally, the integrated feature maps are fed into the detection head to facilitate efficient object detection. Experimental results show that our method has improved the average precision of vehicle detection by 1.01% on the KITTI dataset over three levels of difficulty compared to Voxel-RCNN. In addition, visualization analyses show that our proposed algorithm also exhibits superior performance in object detection.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001245388300001  
Indexed Date:  
2024-06-18  
Institution Address:  
Ocean Univ China, Fac Informat Sci & Engn, Qingdao 266100, Peoples R China  
Institution Address:  
Zhejiang Wanli Univ, Ningbo Key Lab DSP, Ningbo 315100, Peoples R China  
Institution:  
Institution:

### 161. Attention Refined Network for Human Pose Estimation

摘要: Recently, multi-scale feature fusion has been considered as one of the most important issues in designing convolutional neural networks (CNNs). However, most existing methods directly add the corresponding layers together without considering the semantic gaps between them, which may lead to inadequately feature fusion results. In this paper, we propose an attention refined network (HR-ARNet) to enhance multi-scale feature fusion for human pose estimation. The HR-ARNet employs channel and spatial attention mechanisms to reinforce important features and suppress unnecessary ones. To tackle the problem of inconsistent among keypoints, we utilize self-attention strategy to model long-range keypoints dependencies. We also propose to use the focus loss, which modifies the commonly used square error loss function to let it mainly focus on top K 'hard' keypoints during training. Focus loss selects 'hard' keypoints based on the training loss and only backpropagates the gradients from the selected keypoints. Experiments on human pose estimation benchmark, MPII Human Pose Dataset and COCO Keypoint Dataset, show that our method can boost the performance of state-of-the-art human pose estimation networks including HRNet (high-resolution net) (Sun et al., Proceedings of the IEEE conference on computer vision and pattern recognition, 2019). The code and models are available at: http://github/tongjiangwei/ARNet.  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
14  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
15  
ISSN:  
1370-4621  
eISSN:  
1573-773X  
入藏号:  
WOS:000652450500002  
Indexed Date:  
2021-06-08  
Institution Address:  
Shanghai Univ, Sch Commun & Informat Engn, Baoshan, Peoples R China  
Institution:

### 162. Improved SSD model in extraction application of expressway toll station

摘要: The locations of expressway toll stations from GaoFen 2 remote sensing images were extracted as the research object.Expressway toll stations and 0.8 mremote sensing images of Beijing,Shanxi,Henan,Guangdong and Fujian in 2019 were selected to create a training sample dataset via image preprocessing,sample labeling,cropping,data enhancement,and sample dataset partition.Multiscale feature fusion was introduced to improve the target detection model of the single-shot multibox detector(SSD)by adding two operations,namely,deconvolution andconcat.The semantic features of high-level feature maps were assigned to low-level feature maps to enhance the upsampling quality and feature fusion capabilities,thereby improved the detection performance on small targets toll stations.The improved SSD model was applied to extract the locations of toll stations in Fujian in 2019 from GaoFen 2 images.The images were automatically sliced along the Fujian highway network vectors,and the slices were input into the model for target detection.The slices with toll stations were retained,and non-maximum suppression was adopted to remove redundant detection frames.The coordinates of the remaining detection frames were transformed into the coordinates of the center points,and the center point vectors of the expressway toll stations were directly output.Thus,the automatic end-to-end extraction of toll station locations could be realized.Research results show that the accuracy and recall of the improved SSD model and their harmonic average are 0.86,0.88,and 0.87, respectively,which are higher than those of the conventional SSD,VGG,Faster R-CNN,and Feature Pyramid Networks(FPN)models.Therefore,the proposed automatic extraction method for toll station locations can considerably improve management efficiency and adequately satisfy the actual needs of highway managers.  
摘要:  
以高分二号遥感影像中的高速公路收费站为研究对象,选取了北京、山西、河南、广东、福建5个省市2019年的高速公路收费站点位和0.8 m遥感影像,通过图像预处理、样本标注、裁切、数据增强、样本集划分的步骤制作训练样本集;引入多尺度特征融合的方法对SSD目标检测模型进行改进,通过增加转置卷积和拼接操作,将高层次特征图像的语义特征赋予低层次特征图像,以增强上采样质量与特征融合能力,从而提升了模型对小目标收费站的检测效果;将改进SSD模型用于2019年福建省高分二号影像中的收费站点位提取,沿福建省高速公路路网矢量对影像进行自动切片,将切片输入模型中进行目标检测;保留有收费站的切片,使用非极大值抑制去除多余的检测框,将剩余的检测框的坐标变换为中心点的坐标,可以直接输出得到高速公路收费站的中心点矢量,从而实现对于收费站点位的端到端自动化提取。研究结果表明:改进SSD模型的精度、召回率及二者的调和平均数分别为0.86、0.88和0.87,均优于传统的SSD,VGG,Faster R-CNN和特征金字塔网络模型。可见,对收费站点位的自动提取可以大大提高公路管理者的工作效率,有效满足公路管理者的实际工作需求。  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1671-1637  
入藏号:  
CSCD:6984910  
Indexed Date:  
2021-09-03  
Institution Address:  
中国公路工程咨询集团有限公司, 北京 100089, 中国  
Institution Address:  
中国交通通信信息中心, 北京 100011, 中国  
Institution Address:  
China Highway Engineering Consulting Group Co.,Ltd., Beijing 100089, China  
Institution Address:  
China Transport Telecommunications and Information Center, Beijing 100011, China  
Institution:  
中国公路工程咨询集团有限公司  
Institution:  
中国交通通信信息中心  
Institution:  
China Highway Engineering Consulting Group Co.,Ltd.  
Institution:  
China Transport Telecommunications and Information Center

### 163. MHT: A multiscale hourglass-transformer for remaining useful life

摘要: Remaining useful life (RUL) prediction of aircraft engines is significant in the health monitoring, operation, and maintenance of aircraft. Capturing more comprehensive device degradation trends at different time scales and extracting long-term dependencies effectively among elements in long time series are two challenges in the field of aircraft engine RUL estimation. To address the aforementioned challenges, this paper proposes a novel multiscale Hourglass-Transformer (MHT) aircraft engine RUL prognostics. Specifically, an hourglass-shaped multiscale feature extractor (HME) is designed based on one-dimensional convolutional neural network, which can scale the time sequence into multi-time scales for feature fusion. Then, a transformer network is employed to further extract features from the fused feature map and output the RUL. To enhance inter-scale data attention, a pyramid self-attention mechanism is employed in both the encoder and decoder. Finally, the superiority and effectiveness of this approach are verified on the Commercial Modular Aero-Propulsion System Simulation (CMAPSS) dataset. Furthermore, the robustness and generalization capability of this method are further validated on New Commercial Modular Aero-Propulsion System Simulation (N-CMAPSS) dataset.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001125930000001  
Indexed Date:  
2023-12-30  
Institution Address:  
Wuhan Univ Technol, Sch Mech & Elect Engn, Wuhan 430070, Peoples R China  
Institution Address:  
Hubei Digital Mfg Key Lab, Wuhan 430070, Peoples R China  
Institution:  
Institution:  
Hubei Digital Mfg Key Lab

### 164. Spatial Small Target Detection Method Based on Multi-Scale Feature

摘要: Small target detection has become an important part of space exploration missions. The existence of weak illumination and interference from the background of star charts in deep and distant space has brought great challenges to space target detection. In addition, the distance of space targets is usually far, so most of them are small targets in the image, and the detection of small targets is also very difficult. To solve the above problems, we propose a multi-scale feature fusion pyramid network. First, we propose the CST module of a CNN fused with Swin Transformer as the feature extraction module of the feature pyramid network to enhance the extraction of target features. Then, we improve the SE attention mechanism and construct the CSE module to find the attention region in the dense star map background. Finally, we introduce improved spatial pyramid pooling to fuse more features to increase the sensory field to obtain multi-scale object information and improve detection performance for small targets. We provide two versions and conducted a detailed ablation study to empirically validate the effectiveness and efficiency of the design of each component in our network architecture. The experimental results show that our network improved in performance compared to the existing feature pyramid.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001269893800001  
Indexed Date:  
2024-07-23  
Institution Address:  
Chinese Acad Sci, Natl Space Sci Ctr, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution Address:  
Beijing Inst Technol, Sch Mech & Elect, Beijing 100081, Peoples R China  
Institution:  
Institution:  
Institution:

### 165. DMFNet: Dual-Encoder Multistage Feature Fusion Network for Infrared

摘要: Infrared small target detection (IRSTD) is the challenging task of identifying small targets with low signal-to-noise ratios in complex backgrounds. Traditional methods in the complex background of IRSTD lead to a large number of false alarms and misdetections. Although CNN-based methods have made progress in IRSTD, how to extract more effective information and fully utilize interlayer information remains an unresolved issue. Therefore, this article proposed a dual-encoder multistage feature fusion network (DMFNet). Specifically, we designed a dual-encoder with different inputs to capture more effective small target feature information. We then designed a receptive field expansion attention module (REAM) to incorporate nonlocal contextual information. In the decoding phase, the Triple Cross-layer Fusion Module (TCFM) was developed to exchange the low-level spatial details and the high-level semantic information for preserving more small target information in deeper layers. Finally, by concatenating multiscale features from various layers of the decoder, more discriminative feature maps were generated to clearly describe the infrared small targets. Experimental results on the NUDT-SIRST, NUAA-SIRST, and IRSTD-1k datasets demonstrated that DMFNet outperforms some other state-of-the-art methods, achieving superior detection performance. The codes are available at https://github.com/BJZHOU2000/DMFNet.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001193238700017  
Indexed Date:  
2024-04-12  
Institution Address:  
Chongqing Univ Posts & Telecommun, Sch Commun & Informat Engn, Chongqing 400065, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Ctr Appl Math Jiangsu Prov, Nanjing 210044, Peoples R China  
Institution Address:  
Chongqing Univ, Coll Comp Sci, Chongqing 400044, Peoples R China  
Institution Address:  
Chongqing Univ, Sch Microelect & Commun Engn, Chongqing 400044, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Chongqing Key Lab Image Cognit, Chongqing 400065, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 166. A Novel Strategy for Global Lane Detection Based on Key-Point Regression

摘要: Lane detection is a fundamental task for autonomous driving. Most existed deep learning-based methods use a combination of semantic segmentation and post-processing for lane information extraction. Such methods not only tend to ignore global lane information but also bring the problem of low efficiency due to the complex models. To solve these problems, a novel global lane detection method based on key-point regression and multi-scale features fusion (KP-MFF) is proposed in this study. Firstly, a regression strategy is presented to generate key-point sequences in each grid of the image for locating the lane. Moreover, a multi-scale feature fusion module is proposed to merge feature maps of different scales. Additionally, a rule-based fast post-processing method is proposed to deal with the series of key-point sequences output by the CNN model, which further improves the lane detection accuracy. Experiments on CULane and TuSimple datasets demonstrate that the proposed method performs more effectively (417 FPS on NVIDIA 2080Ti and 91 FPS on NVIDIA Jetson AGX Xavier) while maintaining competitive accuracy compared with state-of-the-art methods. The road test also validates the practicability and effectiveness of the proposed method.  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
1524-9050  
eISSN:  
1558-0016  
入藏号:  
WOS:000840480000001  
Indexed Date:  
2022-08-25  
Institution Address:  
South China Univ Technol, Sch Mech & Automot Engn, Guangzhou 510641, Peoples R China  
Institution Address:  
South China Univ Technol, Shien Ming Wu Sch Intelligent Engn, Guangzhou 511442, Peoples R China  
Institution Address:  
Pazhou Lab, Guangzhou 510335, Peoples R China  
Institution:  
Institution:  
Institution:

### 167. Multi-scale feature fusion for histopathological image categorisation in

摘要: Histopathological images of the breast play a crucial role in diagnosing breast cancer and developing effective treatment strategies. However, the complexity of these images poses challenges for traditional convolutional neural networks, which struggle to extract all relevant features. In this paper, we present a novel classification method that addresses this limitation and enhances the accuracy of diagnosing benign and malignant breast tumors. Our approach involves data augmentation techniques, such as staining normalization, image patch generation, and spatial geometry transformation, to expand the training set. We utilize BCMNet (Breast Classification Fusion Multi-Scale Feature Network), which combines VGG16 and CBAM, to fuse spatial features, channel features, and multi-scale features of the input images. We evaluate our method on the BreakHis dataset and achieve average accuracies of 91.91%, 91.14%, 92.65%, and 87.56% for breast histopathology image classification at four different magnifications. Compared to the benchmark network VGG16 without data augmentation, our approach improves the highest accuracy by 3.69% and the average accuracy by 18.47%. Moreover, our proposed method achieves an F1 score of 95.02% and an average recognition accuracy of 92.20% at the patient level. Our approach surpasses previous breast pathology image classification methods, both single and hybrid networks, and establishes a strong theoretical foundation for computer-assisted breast cancer diagnosis.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
2168-1163  
eISSN:  
2168-1171  
入藏号:  
WOS:001018755400001  
Indexed Date:  
2023-07-12  
Institution Address:  
Jiangsu Univ Sci & Technol, Sch Comp, Zhenjiang, Peoples R China  
Institution Address:  
Univ Nevada, Dept Elect & Comp Engn, Las Vegas, NV USA  
Institution Address:  
Jiangsu Univ Sci & Technol, Sch Comp, Zhenjiang 212003, Peoples R China  
Institution:  
Institution:  
Institution:

### 168. MSFANet: Multiscale Fusion Attention Network for Road Segmentation of

摘要: With the development of deep learning and remote sensing technologies in recent years, many semantic segmentation methods based on convolutional neural networks (CNNs) have been applied to road extraction. However, previous deep learning-based road extraction methods primarily used RGB imagery as an input and did not take advantage of the spectral information contained in hyperspectral imagery. These methods can produce discontinuous outputs caused by objects with similar spectral signatures to roads. In addition, the images obtained from different Earth remote sensing sensors may have different spatial resolutions, enhancing the difficulty of the joint analysis. This work proposes the Multiscale Fusion Attention Network (MSFANet) to overcome these problems. Compared to traditional road extraction frameworks, the proposed MSFANet fuses information from different spectra at multiple scales. In MSFANet, multispectral remote sensing data is used as an additional input to the network, in addition to RGB remote sensing data, to obtain richer spectral information. The Cross-source Feature Fusion Module (CFFM) is used to calibrate and fuse spectral features at different scales, reducing the impact of noise and redundant features from different inputs. The Multiscale Semantic Aggregation Decoder (MSAD) fuses multiscale features and global context information from the upsampling process layer by layer, reducing information loss during the multiscale feature fusion. The proposed MSFANet network was applied to the SpaceNet dataset and self-annotated images from Chongzhou, a representative city in China. Our MSFANet performs better over the baseline HRNet by a large margin of +6.38 IoU and +5.11 F1-score on the SpaceNet dataset, +3.61 IoU and +2.32 F1-score on the self-annotated dataset (Chongzhou dataset). Moreover, the effectiveness of MSFANet was also proven by comparative experiments with other studies.  
Web of Science 核心合集中的 "被引频次":  
10  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
10  
eISSN:  
2072-4292  
入藏号:  
WOS:000977384700001  
Indexed Date:  
2023-05-14  
Institution Address:  
Univ Elect Sci & Technol China, Sch Automat Engn, Chengdu 611731, Peoples R China  
Institution Address:  
Chengdu Univ Informat Technol, Sch Software Engn, Chengdu, Peoples R China  
Institution:  
Institution:

### 169. Subpixel Multilevel Scale Feature Learning and Adaptive Attention

摘要: Convolutional neural networks (CNNs) play an important role in hyperspectral image (HSI) classification due to their powerful feature extraction ability. Multiscale information is an important means of enhancing the feature representation ability. However, current HSI classification models based on deep learning only use fixed patches as the network input, which may not well reflect the complexity and richness of HSIs. While the existing methods achieve good classification performance for large-scale scenes, the classification of boundary locations and small-scale scenes is still challenging. In addition, dimensional dislocation often exists in the feature fusion process, and the up/downsampling operation for feature alignment may introduce extra noise or result in feature loss. Aiming at the above issues, this paper deeply explores multiscale features, proposes an adaptive attention constraint fusion module for different scale features, and designs a semantic feature enhancement module for high-dimensional features. First, HSI data of two different spatial scales are fed into the model. For the two inputs, we upsample them using bilinear interpolation to obtain their subpixel data. The proposed multiscale feature extraction module is intended to extract the features of the above four parts of the data. For the extracted features, the multiscale attention fusion module is used for feature fusion, and then, the fused features are fed into the high-level feature semantic enhancement module. Finally, based on the fully connected layer and softmax layer, the prediction results of the proposed model are obtained. Experimental results on four public HSI databases verify that the proposed method outperforms several state-of-the-art methods.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
eISSN:  
2072-4292  
入藏号:  
WOS:000839735600001  
Indexed Date:  
2022-08-23  
Institution Address:  
Nanjing Univ Sci & Technol, Sch Comp Sci & Engn, Nanjing 210094, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Jiangsu Key Lab Broadband Wireless Commun & Inter, Nanjing 210003, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Sch Internet Things, Nanjing 210003, Peoples R China  
Institution:  
Institution:  
Institution:

### 170. A Hybrid CNN-Transformer Network for Object Detection in Optical Remote

摘要: Remote sensing images (RSIs) object detection is important in natural disaster management, urban planning and resource exploration. However, due to the large differences between RSIs and natural images (NIs), most of the existing object detectors for NIs cannot be directly used to process RSIs. Most existing models based on convolutional neural networks (CNNs) require additional design of specific attentional modules to relate small targets in RSIs to global positional relationships. In contrast, transformer-based models had to add modules to obtain more detailed information. This imposes additional computational overheads for deployment on edge devices. To solve the above-mentioned problem, we propose a hybrid CNN and transformer model (DConvTrans-LKA) to enhance the model's ability to acquire features and design a fusion of local and global attention mechanisms to fuse local features and global location information. To better fuse the feature and location information extracted by the model, we introduce a feature residual pyramid network to enhance the model's ability to fuse multiscale feature maps. Finally, we conduct experiments in three representative optical RSI datasets (NWPU VHR-10, HRRSD, and DIOR) to verify the effectiveness of our proposed DConvTrans-LKA method. The experimental results show that our proposed method reaches 61.7%, 82.1%, and 61.3% at mAP at 0.5, respectively, further demonstrating the potential of our proposed method in RSI object detection tasks.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001409622100002  
Indexed Date:  
2025-02-09  
Institution Address:  
Nanjing Univ Posts & Telecommun, Coll Telecommun & Informat Engn, Nanjing 210003, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Coll Internet Things, Nanjing 210003, Peoples R China  
Institution Address:  
Hangzhou Dianzi Univ, Coll Commun Engn, Hangzhou 310018, Peoples R China  
Institution:  
Institution:  
Institution:

### 171. Gearbox fault diagnosis based on RGT-MFFIN and multi-sensor fusion image

摘要: In gearbox fault diagnosis based on vibration and torque state data, traditional one-dimensional time-frequency domain analysis methods often suffer from insufficient feature expression and mining, and require complex noise reduction and filtering preprocessing. To address this issue, this paper proposes a fusion image generation method that integrates the advantages of recurrence plot (RP) and Gramian angular summation field (GASF) to generate recurrence Gramian transformed (RGT) images. This approach integrates both global and local fault information, making the fault characteristics more intuitive and easier to analyze. Given that multi-sensor collaboration can enhance feature representation, feature-level fusion increases the computational burden, and decision-level fusion is prone to losing inter-sensor correlation information, this paper adopts data-level fusion for image sample enhancement. In the diagnostic method, the challenge of traditional convolutional neural networks (CNNs) in extracting diverse geometric linear structures from fused images is addressed by introducing deformable convolutional blocks for initial feature extraction. Additionally, a multi-scale feature fusion interaction network (MFFIN) is constructed. This network incorporates a channel-space interactive attention mechanism on top of multi-scale feature extraction, assigning weights to features according to their importance while facilitating the interaction of feature information. Finally, validation is carried out using public datasets, and the experimental results show that the proposed method demonstrates significant advantages in classification accuracy and robustness under variable operating conditions and noise, thereby proving its effectiveness and practicality.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2631-8695  
入藏号:  
WOS:001322178900001  
Indexed Date:  
2024-10-05  
Institution Address:  
Ningbo Univ, Fac Mech Engn & Mech, Ningbo 315211, Peoples R China  
Institution:

### 172. Remote Sensing Image Segmentation Network Based on Adaptive Multiscale

摘要: Remote sensing image segmentation algorithms are susceptible to interference from environmental factors, such as object occlusion and uneven illumination. Existing deep learning remote sensing image semantic segmentation methods usually adopt an end-to-end codec structure. However, they still suffer from inaccurate segmentation for the structure and contours of high similarity objects. Therefore, to improve the algorithm robustness and classification accuracy, a deep convolutional neural network remote sensing image semantic segmentation algorithm based on contour gradient learning is proposed. To improve the quality of the predicted feature maps, the adaptive attention-based multichannel multiscale feature fusion network (D-MMA Net) is proposed based on the SegNet model network. The D-MA block uses an attention-based adaptive multiscale module to adaptively extract different scale features according to the learned weights to obtain more effective high level semantic features. To further refine the extracted object boundaries, the contour extraction module, a learnable contour extraction module, is proposed based on the principle of the Sobel edge detection operator. Finally, the contour information is combined with multi-scale semantic features to enhance the robustness of the spatial resolution of the image. The experimental results show that the proposed method improves the segmentation accuracy and produces good segmentation results for irregular object boundaries.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1006-4125  
入藏号:  
WOS:001154750000044  
Indexed Date:  
2024-03-13  
Institution Address:  
Guizhou Univ, Coll Comp Sci & Technol, Guiyang 550025, Guizhou, Peoples R China  
Institution Address:  
Guiyang Orbita Aerosp Sci & Technol Co Ltd, Guiyang 550027, Guizhou, Peoples R China  
Institution Address:  
Guizhou Educ Univ, Big Data Sci & Intelligent Engn Res Inst, Guiyang 550018, Guizhou, Peoples R China  
Institution:  
Institution:  
Guiyang Orbita Aerosp Sci & Technol Co Ltd  
Institution:

### 173. Research on Land Use and Land Cover Information Extraction Methods for

摘要: To address the challenges that convolutional neural networks (CNNs) face in extracting small objects and handling class imbalance in remote sensing imagery, this paper proposes a novel spatial contextual information and multiscale feature fusion encoding-decoding network, SCIMF-Net. Firstly, SCIMF-Net employs an improved ResNeXt-101 deep backbone network, significantly enhancing the extraction capability of small object features. Next, a novel PMFF module is designed to effectively promote the fusion of features at different scales, deepening the model's understanding of global and local spatial contextual information. Finally, introducing a weighted joint loss function improves the SCIMF-Net model's performance in extracting LULC information under class imbalance conditions. Experimental results show that compared to other CNNs such as Res-FCN, U-Net, SE-U-Net, and U-Net++, SCIMF-Net improves PA by 0.68%, 0.54%, 1.61%, and 3.39%, respectively; MPA by 2.96%, 4.51%, 2.37%, and 3.45%, respectively; and MIOU by 3.27%, 4.89%, 4.2%, and 5.68%, respectively. Detailed comparisons of locally visualized LULC information extraction results indicate that SCIMF-Net can accurately extract information from imbalanced classes and small objects.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2220-9964  
入藏号:  
WOS:001365425200001  
Indexed Date:  
2024-12-05  
Institution Address:  
Yunnan Normal Univ, Fac Geog, Kunming 650500, Peoples R China  
Institution Address:  
Key Lab Resources & Remote Sensing Univ Yunnan Kun, Kunming 650500, Peoples R China  
Institution Address:  
Ctr Geospatial Informat Engn & Technol Yunnan Prov, Kunming 650500, Peoples R China  
Institution Address:  
Yunnan Normal Univ, Sch Informat Sci & Technol, Kunming 650500, Peoples R China  
Institution:  
Institution:  
Key Lab Resources & Remote Sensing Univ Yunnan Kun  
Institution:  
Ctr Geospatial Informat Engn & Technol Yunnan Prov  
Institution:

### 174. Compound Multiscale Weak Dense Network with Hybrid Attention for

摘要: Recently, hyperspectral image (HSI) classification has become a popular research direction in remote sensing. The emergence of convolutional neural networks (CNNs) has greatly promoted the development of this field and demonstrated excellent classification performance. However, due to the particularity of HSIs, redundant information and limited samples pose huge challenges for extracting strong discriminative features. In addition, addressing how to fully mine the internal correlation of the data or features based on the existing model is also crucial in improving classification performance. To overcome the above limitations, this work presents a strong feature extraction neural network with an attention mechanism. Firstly, the original HSI is weighted by means of the hybrid spectral-spatial attention mechanism. Then, the data are input into a spectral feature extraction branch and a spatial feature extraction branch, composed of multiscale feature extraction modules and weak dense feature extraction modules, to extract high-level semantic features. These two features are compressed and fused using the global average pooling and concat approaches. Finally, the classification results are obtained by using two fully connected layers and one Softmax layer. A performance comparison shows the enhanced classification performance of the proposed model compared to the current state of the art on three public datasets.  
Web of Science 核心合集中的 "被引频次":  
5  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
1  
被引频次合计:  
5  
eISSN:  
2072-4292  
入藏号:  
WOS:000689827000001  
Indexed Date:  
2021-09-06  
Institution Address:  
Nanjing Univ Sci & Technol, Sch Comp Sci & Engn, Nanjing 210094, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Jiangsu Key Lab Broadband Wireless Commun & Inter, Nanjing 210003, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Sch Internet Things, Nanjing 210003, Peoples R China  
Institution:  
Institution:  
Institution:

### 175. DMs-MAFM plus EfficientNet: a hybrid model for predicting dysthyroid

摘要: Thyroid-associated ophthalmopathy (TAO) is a very common autoimmune orbital disease. Approximately 4%-8% of TAO patients will deteriorate and develop the most severe dysthyroid optic neuropathy (DON). According to the current data provided by clinical experts, there is still a certain proportion of suspected DON patients who cannot be diagnosed, and the clinical evaluation has low sensitivity and specificity. There is an urgent need for an efficient and accurate method to assist physicians in identifying DON. This study proposes a hybrid deep learning model to accurately identify suspected DON patients using computed tomography (CT). The hybrid model is mainly composed of the double multiscale and multi attention fusion module (DMs-MAFM) and a deep convolutional neural network. The DMs-MAFM is the feature extraction module proposed in this study, and it contains a multiscale feature fusion algorithm and improved channel attention and spatial attention, which can capture the features of tiny objects in the images. Multiscale feature fusion is combined with an attention mechanism to form a multilevel feature extraction module. The multiscale fusion algorithm can aggregate different receptive field features, and then fully obtain the channel and spatial correlation of the feature map through the multiscale channel attention aggregation module and spatial attention module, respectively. According to the experimental results, the hybrid model proposed in this study can accurately identify suspected DON patients, with Accuracy reaching 96%, Specificity reaching 99.5%, Sensitivity reaching 94%, Precision reaching 98.9% and Fl-score reaching 96.4%. According to the evaluation by experts, the hybrid model proposed in this study has some enlightening significance for the diagnosis and prediction of clinically suspect DON.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
8  
ISSN:  
0140-0118  
eISSN:  
1741-0444  
入藏号:  
WOS:000855996900001  
PubMed ID:  
36129645  
Indexed Date:  
2022-09-28  
Institution Address:  
Hubei Univ Technol, Sch Comp Sci, Nanli St 28, Wuhan 430068, Peoples R China  
Institution Address:  
Huazhong Univ Sci & Technol, Tongji Med Coll, Union Hosp, Zhongshan Pk, Wuhan 430022, Peoples R China  
Institution:  
Institution:

### 176. Hidden Feature-Guided Semantic Segmentation Network for Remote Sensing

摘要: For semantic segmentation of remote sensing images, convolutional neural networks (CNNs) have proven to be powerful tools. However, the existing CNN-based methods have the problems of feature information loss, serious interference by clutter information, and ignoring the correlation between different scale features. To solve these problems, this article proposes a novel hidden feature-guided semantic segmentation network (HFGNet) for remote sensing images, which achieves accurate semantic segmentation by hierarchically extracting and fusing valuable feature information. Specifically, the hidden feature extraction module (HFE-M) is introduced to suppress the salient feature representation to mine more valuable hidden features. Meanwhile, the multifeature interactive fusion module (MIF-M) establishes the correlation between different features to achieve hierarchical feature fusion. The multiscale feature calibration module (MSFC) is constructed to enhance the diversity and refinement representation of hierarchical fusion features. Besides, the local-channel attention mechanism (LCA-M) is designed to improve the feature perception capability of the object region and suppress background information interference. We conducted extensive experiments on the widely used ISPRS 2-D Semantic Labeling dataset and the 15-Class Gaofen Image dataset. Experimental results demonstrate that the proposed HFGNet has advantages over several state-of-the-art methods. The source code and models are available at https://github.com/darkseid-arch/RS-HFGNet.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
21  
被引频次合计:  
22  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000967003300001  
Indexed Date:  
2023-04-30  
Institution Address:  
Xijing Univ, Sch Elect Informat, Xian 710123, Peoples R China  
Institution Address:  
Tianjin Univ Sci & Technol, Sch Artificial Intelligence, Tianjin 300457, Peoples R China  
Institution Address:  
AF Engn Univ, Sch Informat & Nav, Xian 710082, Peoples R China  
Institution:  
Institution:  
Institution:

### 177. FI-Net: Rethinking Feature Interactions for Medical Image Segmentation

摘要: To solve the problems of existing hybrid networks based on convolutional neural networks (CNN) and Transformers, we propose a new encoder-decoder network FI-Net based on CNN-Transformer for medical image segmentation. In the encoder part, a dual-stream encoder is used to capture local details and long-range dependencies. Moreover, the attentional feature fusion module is used to perform interactive feature fusion of dual-branch features, maximizing the retention of local details and global semantic information in medical images. At the same time, the multi-scale feature aggregation module is used to aggregate local information and capture multi-scale context to mine more semantic details. The multi-level feature bridging module is used in skip connections to bridge multi-level features and mask information to assist multi-scale feature interaction. Experimental results on seven public medical image datasets fully demonstrate the effectiveness and advancement of our method. In future work, we plan to extend FI-Net to support 3D medical image segmentation tasks and combine self-supervised learning and knowledge distillation to alleviate the overfitting problem of limited data training.  
A new encoder-decoder network FI-Net based on convolutional neural networks (CNN)-transformer is proposed for medical image segmentation. It rethinks the uniqueness of feature interactions in medical images to design four effective modules. Experimental results on seven public medical image datasets fully demonstrate the effectiveness and advancement of the method.image (c) 2024 WILEY-VCH GmbH  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2640-4567  
入藏号:  
WOS:001223873600001  
Indexed Date:  
2024-05-23  
Institution Address:  
Cent South Univ, Sch Comp Sci & Engn, Changsha 410000, Peoples R China  
Institution Address:  
Cent South Univ, Xiangya Hosp, Dept Urol, Changsha 410008, Peoples R China  
Institution Address:  
Cent South Univ, Xiangya Hosp, Dept Burns & Plast Surg, Changsha 410008, Peoples R China  
Institution Address:  
Cent South Univ, Xiangya Hosp, Natl Clin Res Ctr Geriatr Disorders, Changsha 410008, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 178. Buried target detection method for ground penetrating radar based on

摘要: Deep learning method has been extensively applied to ground penetrating radar two-dimensional profile (GPR B-SCAN) hyperbola detection recently. We propose a B-SCAN image feature extraction method based on the constraints of the GPR physical model, and further detect the weak boundary feature curve of the target in the local space. A deep convolutional neural network (DCNN) is first designed to extract high-level semantic features from B-SCAN images to remove direct wave. Next, a multiscale feature fusion DCNN is used to extract the features of the B-SCAN image with the direct wave removed, and the classifier network is used to identify the hyperbola of the upper boundary feature of the target. Finally, according to the hyperbola, the local space corresponding to the target in the B-SCAN image is determined. On this basis, the amplitude and phase information of the scattered electric field are used to segment the lower boundary characteristic curve of the target through convolution operation. Experimental results on simulation and field data show that feature information of the buried target in the GPR B-SCAN image can be efficiently extracted when the proposed method is adopted. (C) The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
10  
eISSN:  
1931-3195  
入藏号:  
WOS:000777198700064  
Indexed Date:  
2022-04-28  
Institution Address:  
Guilin Univ Elect Technol, Sch Informat & Commun, Guilin, Peoples R China  
Institution Address:  
Hezhou Univ, Sch Artificial Intelligence, Hezhou, Peoples R China  
Institution Address:  
Guilin Univ Elect Technol, Satellite Nav Positioning & Locat Serv Natl & Loc, Guilin, Peoples R China  
Institution Address:  
Shanxi Transportat Technol R&D Co Ltd, Taiyuan, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Shanxi Transportat Technol R&D Co Ltd

### 179. M2F2-RCNN: Multi-functional Faster RCNN Based on Multi-scale Feature

摘要: In order to realize fast and accurate search of sensitive regions in remote sensing images, we propose a multi-functional faster RCNN based on multi-scale feature fusion model for region search. The feature extraction network is based on ResNet50 and the dilated residual blocks are utilized for multi-layer and multi-scale feature fusion. We add a path aggregation network with a convolution block attention module (CBAM) attention mechanism in the backbone network to improve the efficiency of feature extraction. Then, the extracted feature map is processed, and RoIAlign is used to improve the pooling operation of regions of interest and it can improve the calculation speed. In the classification stage, an improved non maximum suppression is used to improve the classification accuracy of the sensitive region. Finally, we conduct cross validation experiments on Google Earth dataset and the DOTA dataset. Meanwhile, the comparison experiments with the state-ofthe-art methods also prove the high efficiency of the proposed method in region search ability.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
1820-0214  
入藏号:  
WOS:001083912900004  
Indexed Date:  
2023-11-11  
Institution Address:  
Harbin Engn Univ, Coll Informat & Commun Engn, Harbin 150001, Peoples R China  
Institution Address:  
Dalian Minzu Univ, Coll Informat & Commun Engn, Dalian 116000, Peoples R China  
Institution Address:  
Tongji Univ, Coll Surveying & Geoinformat, Shanghai 200092, Peoples R China  
Institution Address:  
Univ Novi Sad, Fac Sci, Novi Sad 21000, Serbia  
Institution Address:  
China Univ Geosci, Sch Informat Engn, Beijing 100083, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 180. Multi-scale feature fusion based attribute-aware face image

摘要: Convolutional neural networks(CNN)have achieved many satisfactory results in face image super-resolution tasks with its powerful representation ability. However,most of the CNN-based methods adopt an increasing network depth approach which demands more storage space and are less applicable in practice. In order to solve this problem, this paper proposes a multi-scale feature fusion based attribute-aware face image super-resolution network. First, this paper adopts a local residual learning via shortcut connection and element-wise addition to reduce the complexity and distill the representative multi-scale features. Second, this paper adopts an attribute-aware module to adaptively rescale and fuse multi-scale features and facial prior which provide rich semantic information to help better super-resolve face images. Third,the network proposed here is constructed by a set of sub-modules, in which the features from each sub-module can be jointly learned by using a hierarchical feature fusion module. Extensive experiments show that the proposed network achieves more appealing results than the state-of-the-art methods; moreover, the method not only generates perceptually plausible face images,but also is able to manipulate the results by adjusting the attribute information.  
摘要:  
卷积神经网络凭借其强大的表征能力,在图像超分辨率任务上取得了许多令人满意的结果。许多基于神经网络的方法采用增加网络深度的方式,存在存储空间消耗多、实用性不强的问题。为解决该问题,该文提出一种基于多尺度特征融合的属性感知人脸图像超分辨率网络。该文借助局部残差模块和逐元素相加的融合方式以减少网络复杂性并提炼出表征能力优秀的多尺度特征。该文构建一个可自适应地融合多尺度特征和人脸先验的属性感知模块,使得网络学习到更丰富的语义信息。该文提出的网络由多个网络子模块级联构成,并通过一个多层次特征融合模块进行共同学习。试验表明:该文方法能取得良好的超分辨率性能,输出更加真实的人脸图像,可以通过调整人脸属性信息进行人脸图像生成效果的操纵。  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1005-9830  
入藏号:  
CSCD:7332684  
Indexed Date:  
2023-02-25  
Institution Address:  
南京理工大学计算机科学与工程学院, 南京, 江苏 210094, 中国  
Institution Address:  
School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China  
Institution:  
南京理工大学计算机科学与工程学院  
Institution:  
School of Computer Science and Engineering, Nanjing University of Science and Technology

### 181. ResLMFFNet: a real-time semantic segmentation network for precision

摘要: Lightweight multiscale-feature-fusion network (LMFFNet), a proficient real-time CNN architecture, adeptly achieves a balance between inference time and accuracy. Capturing the intricate details of precision agriculture target objects in remote sensing images requires deep SEM-B blocks in the LMFFNet model design. However, employing numerous SEM-B units leads to instability during backward gradient flow. This work proposes the novel residual-LMFFNet (ResLMFFNet) model for ensuring smooth gradient flow within SEM-B blocks. By incorporating residual connections, ResLMFFNet achieves improved accuracy without affecting the inference speed and the number of trainable parameters. The results of the experiments demonstrate that this architecture has achieved superior performance compared to other real-time architectures across diverse precision agriculture applications involving UAV and satellite images. Compared to LMFFNet, the ResLMFFNet architecture enhances the Jaccard Index values by 2.1% for tree detection, 1.4% for crop detection, and 11.2% for wheat-yellow rust detection. Achieving these remarkable accuracy levels involves maintaining almost identical inference time and computational complexity as the LMFFNet model. The source code is available on GitHub: https://github.com/iremulku/Semantic-Segmentation-in-Precision-Agriculture.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001233962300002  
Indexed Date:  
2024-06-03  
Institution Address:  
Ankara Univ, Dept Comp Engn, TR-06830 Ankara, Turkiye  
Institution:

### 182. Multi-Scale Feature Fusion Convolutional Neural Network for Indoor Small

摘要: The development of object detection technology makes it possible for robots to interact with people and the environment, but the changeable application scenarios make the detection accuracy of small and medium objects in the practical application of object detection technology low. In this paper, based on multi-scale feature fusion of indoor small target detection method, using the device to collect different indoor images with angle, light, and shade conditions, and use the image enhancement technology to set up and amplify a date set, with indoor scenarios and the SSD algorithm in target detection layer and its adjacent features fusion. The Faster R-CNN, YOLOv5, SSD, and SSD target detection models based on multi-scale feature fusion were trained on an indoor scene data set based on transfer learning. The experimental results show that multi-scale feature fusion can improve the detection accuracy of all kinds of objects, especially for objects with a relatively small scale. In addition, although the detection speed of the improved SSD algorithm decreases, it is faster than the Faster R-CNN, which better achieves the balance between target detection accuracy and speed.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
90  
在 BIOSIS Citation Index 中的被引频次:  
22  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
91  
ISSN:  
1662-5218  
入藏号:  
WOS:000806199800001  
PubMed ID:  
35663726  
Indexed Date:  
2022-06-09  
Institution Address:  
Wuhan Univ Sci & Technol, Coll Comp Sci & Technol, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Hubei Prov Key Lab Intelligent Informat Proc & Rea, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Key Lab Met Equipment, Control Technol Minist Educ, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Hubei Key Lab Mech Transmiss & Mfg Engn, Wuhan, Peoples R China  
Institution Address:  
Wuhan Univ Sci & Technol, Precis Mfg Res Inst, Wuhan, Peoples R China  
Institution Address:  
Univ Portsmouth, Sch Creat Technol, Portsmouth, England  
Institution Address:  
Univ Portsmouth, Sch Energy & Elect Engn, Portsmouth, England  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 183. PSMFNet: Lightweight Partial Separation and Multiscale Fusion Network

摘要: The employment of deep convolutional neural networks has recently contributed to significant progress in single image super-resolution (SISR) research. However, the high computational demands of most SR techniques hinder their applicability to edge devices, despite their satisfactory reconstruction performance. These methods commonly use standard convolutions, which increase the convolutional operation cost of the model. In this paper, a lightweight Partial Separation and Multiscale Fusion Network (PSMFNet) is proposed to alleviate this problem. Specifically, this paper introduces partial convolution (PConv), which reduces the redundant convolution operations throughout the model by separating some of the features of an image while retaining features useful for image reconstruction. Additionally, it is worth noting that the existing methods have not fully utilized the rich feature information, leading to information loss, which reduces the ability to learn feature representations. Inspired by self-attention, this paper develops a multiscale feature fusion block (MFFB), which can better utilize the non-local features of an image. MFFB can learn long-range dependencies from the spatial dimension and extract features from the channel dimension, thereby obtaining more comprehensive and rich feature information. As the role of the MFFB is to capture rich global features, this paper further introduces an efficient inverted residual block (EIRB) to supplement the local feature extraction ability of PSMFNet. A comprehensive analysis of the experimental results shows that PSMFNet maintains a better performance with fewer parameters than the state-of-the-art models.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001352657700001  
Indexed Date:  
2024-11-17  
Institution Address:  
GDAS, Guangdong Key Lab Modern Control Technol, Inst Intelligent Mfg, Guangzhou 510030, Peoples R China  
Institution Address:  
South China Univ Technol, Sch Mech & Automot Engn, Guangzhou 511442, Peoples R China  
Institution Address:  
Kunming Univ Sci & Technol, Fac Informat Engn & Automat, Kunming 650500, Peoples R China  
Institution Address:  
Wuyi Univ, Sch Fac Intelligent Mfg, Jiangmen 529020, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 184. Automatic Detection of Electric Motorcycle Based on Improved YOLOv5s

摘要: Electric motorcycles are widely used due to their economic, portable, and easy-to-use characteristics. Power batteries are the primary power source of electric motorcycles. Electric motorcycles are usually pushed into elevators and parked at home or in enclosed corridor spaces for charging, which may pose serious safety hazards due to using inferior or expired batteries. The traditional manual management method is limited by human resources, making it difficult to manage and monitor such behavior. Automated detection of electric motorcycles based on artificial intelligence technology is an effective solution. Considering that common monitoring systems typically have limited data processing capabilities, this study proposes an electric motorcycle detection model based on improved You Only Look Once version 5s (YOLOv5s). Firstly, we develop the model by adding a transformer encoder module to the backbone of classical YOLOv5s. Next, the Bidirectional Feature Pyramid Network (BiFPN) is used for cross-scale connectivity and multiscale feature fusion. Finally, the Coordinate Attention module (CA) is added to improve the representation capacity of the target features and enhance the detection accuracy. The results of comparative experiments and ablation experiments verified the effective performance of the proposed model, which attained a mean average precision of 81.2%. Compared to classical models like faster R-CNN and YOLOv5, this methodology achieves higher performance with fewer parameters and computational complexity, meeting real-time requirements.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2090-0147  
eISSN:  
2090-0155  
入藏号:  
WOS:001288286400001  
Indexed Date:  
2024-08-16  
Institution Address:  
Jimei Univ, Chengyi Coll, Xiamen 363000, Fujian, Peoples R China  
Institution:

### 185. ECF-Net: Enhanced, Channel-Based, Multi-Scale Feature Fusion Network for

摘要: Accurate segmentation of COVID-19 lesion regions in lung CT images aids physicians in analyzing and diagnosing patients' conditions. However, the varying morphology and blurred contours of these regions make this task complex and challenging. Existing methods utilizing Transformer architecture lack attention to local features, leading to the loss of detailed information in tiny lesion regions. To address these issues, we propose a multi-scale feature fusion network, ECF-Net, based on channel enhancement. Specifically, we leverage the learning capabilities of both CNN and Transformer architectures to design parallel channel extraction blocks in three different ways, effectively capturing diverse lesion features. Additionally, to minimize irrelevant information in the high-dimensional feature space and focus the network on useful and critical information, we develop adaptive feature generation blocks. Lastly, a bidirectional pyramid-structured feature fusion approach is introduced to integrate features at different levels, enhancing the diversity of feature representations and improving segmentation accuracy for lesions of various scales. The proposed method is tested on four COVID-19 datasets, demonstrating mIoU values of 84.36%, 87.15%, 83.73%, and 75.58%, respectively, outperforming several current state-of-the-art methods and exhibiting excellent segmentation performance. These findings provide robust technical support for medical image segmentation in clinical practice.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2079-9292  
入藏号:  
WOS:001311217200001  
Indexed Date:  
2024-09-21  
Institution Address:  
Gansu Agr Univ, Coll Informat Sci & Technol, Lanzhou 730070, Peoples R China  
Institution Address:  
Ningbo Univ Technol, Sch Cyber Sci & Engn, Ningbo 315211, Peoples R China  
Institution Address:  
Ningbo Inst Digital Twin, Ningbo 315201, Peoples R China  
Institution Address:  
Ningbo 2 Hosp, Ningbo 315010, Peoples R China  
Institution:  
Institution:  
Institution:  
Ningbo Inst Digital Twin  
Institution:  
Ningbo 2 Hosp

### 186. Detection of bergamot diseases and pests based on improved SSD

摘要: [Objectives]Aiming at the difficulty of identifying pests and diseases in the real environment of bergamot in the field,this paper proposed a method for detecting pests and diseases of bergamot based on improved SSD(single shot MultiBox detector) algorithm-SSD-Res50-3C algorithm. [Methods] The backbone network partially replaced the original VGG16 network with the ResNet50 network,and enhanced the model's ability to extract the characteristics of pests and diseases of bergamot in the real background of the field. A lightweight and efficient feature fusion module was added before predicting the feature layer to improve the multi-scale feature fusion capability of the SSD algorithm,and further improved the anti-interference ability of the SSD algorithm in the real background of the field. [Results]The mean average precision of SSD-Res50-3C algorithm was 92.86%,which was 6.61% higher than that of the original SSD algorithm,and the FPS(frames per second)reached 64.1. Compared with YOLO v3,YOLO v4,YOLO v5x6,Faster R-CNN,and EfficientDet-D3 models,the mean average precision of SSD-Res50-3C algorithm was 6.41%,2.01%,0.79%, 0.58%,and 5.10%,respectively,and FPS was 16.20,40.28,24.40,36.20,and 54.84,respectively. [Conclusions]The pests and diseases detection method of bergamot based on the improved SSD algorithm proposed in this paper could weaken the interference information of the real environment in the field,accurately identify the target of pests and diseases of bergamot in the real environment of the field,and provide a new idea for the detection of pests and diseases of bergamot in the real environment of the field.  
摘要:  
[目的]针对广佛手在田间真实环境下病虫害识别较为困难的问题,提出一种基于改进SSD(single shot MultiBox detector)算法SSD-Res50-3C的广佛手病虫害检测方法。[方法]SSD-Res50-3C算法主干网络部分用ResNet50网络替换原有的VGG16网络,增加模型在田间真实环境下对广佛手病虫害特征的提取能力;在预测特征层之前加入一种轻量高效的特征融合模块提升SSD算法的多尺度特征融合能力,进一步提高SSD算法在田间真实环境下的抗干扰能力。[结果]SSDRes50-3C算法平均精度均值达到92.86%,相较原始的SSD算法提升6.61%,FPS(frames per second)达到64.1。相比YOLO v3、 YOLO v4、YOLO v5x6、Faster R-CNN和EfficientDet-D3模型,SSD-Res50-3C算法的平均精度均值分别高6.41%、2.01%、0.79%、 0.58%和5.10%,FPS分别高16.20、40.280、24.40、36.20和54.84。[结论]基于改进SSD算法的广佛手病虫害检测方法能够弱化田间真实环境的干扰信息,能准确识别田间真实环境下广佛手病虫害目标,可为田间真实环境下广佛手病虫害检测提供一种新思路。  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1000-2030  
入藏号:  
CSCD:7540043  
Indexed Date:  
2023-12-01  
Institution Address:  
华南农业大学电子工程学院/人工智能学院, 广州, 广东 510642, 中国  
Institution Address:  
广东省农业技术推广中心, 广州, 广东 510520, 中国  
Institution Address:  
College of Electronic Engineering/College of Artificial Intelligence, South China Agricultural University, Guangzhou, Guangdong 510642, China  
Institution Address:  
Guangdong Agricultural Technology Extension Center, Guangzhou, Guangdong 510520, China  
Institution:  
华南农业大学电子工程学院/人工智能学院  
Institution:  
广东省农业技术推广中心  
Institution:  
College of Electronic Engineering/College of Artificial Intelligence, South China Agricultural University  
Institution:  
Guangdong Agricultural Technology Extension Center

### 187. Impact Load Localization Based on Multi-Scale Feature Fusion

摘要: In order to achieve impact load localization of complex structures such as ships, this paper proposes a multi-scale feature fusion convolutional neural network (MSFF-CNN) method for impact load localization. An end-to-end machine learning model is used, where the raw vibration signals of impact loads are directly fed into the network model to avoid the process of feature extraction. Automatic feature learning and feature concatenation of the signal are achieved through four independent convolutional layers, each using a different size of convolutional kernel. Data normalization and L2 regularization techniques are introduced to enhance the data and prevent overfitting. Classification and localization of impact loads are accomplished using a softmax classification layer. Validation experiments are carried out using a ship's stern compartment model. Our results show that the classification and localization accuracy of the impact load sample group of MSFF-CNN reaches 94.29% compared with a traditional CNN. The method further improves the ability of the network to extract state features, takes local perception and global vision into account, effectively improves the classification ability of the model, and has good prospects for engineering applications.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:001323989600001  
PubMed ID:  
39338807  
Indexed Date:  
2024-10-07  
Institution Address:  
Naval Univ Engn, Lab Vibrat & Noise, Wuhan 430033, Peoples R China  
Institution Address:  
Naval Univ Engn, Natl Key Lab Vibrat & Noise Ship, Wuhan 430033, Peoples R China  
Institution:  
Institution:

### 188. Power profiling analysis method based on multi-scale feature fusion

摘要: In the wave of digitization,the rapid development of 5G and 6G technologies is leading the mobile communication systems into a new era.Advanced hardware devices and encryption chips offer robust support for the escalating demand in data processing and the growing emphasis on security.In this context,various hardware devices equipped with modern cryptographic technology are gradually evolving into indispensable cornerstones of our daily lives.These devices have the capability to resist traditional cryptographic analysis.In recent years,one of the focuses of academic research is the analysis of physical leakage occurring during the actual operation of devices,a field known as Side-Channel Analysis(SCA).Deep learning-driven side-channel analysis has been widely recognized as an effective method.Aiming at the current neural network model's problems such as high demand for the number of traces,poor robustness,and slow convergence speed,this paper proposes a multiscale feature fusion side-channel analysis method based on CNNbest.Firstly,the structure of the feature extraction network is revised to mitigate the issue of deep feature vectors being susceptible to excessive interpretation of noise details and model overfitting.Subsequently,a filtering array is used to perform Discrete Wavelet Transform(DWT) analysis,constructing multi-resolution time-frequency representations to enhance data quality.Finally,a lightweight Convolutional Block Attention Module(CBAM) incorporating channel spatial attention is introduced to improve the learning efficiency of key features in power consumption curves.Experimental results demonstrate that the proposed method reduces the power consumption curves required for side-channel analysis by 88.27% compared to the original model,significantly improving analysis performance and meeting the requirements of side-channel modeling and analysis.  
摘要:  
在数字化浪潮推动下,5G和6G技术的快速发展正引领移动通信系统步入新阶段.先进的硬件设备和加密芯片为不断增长的数据处理需求和日益关注的安全保障提供了强有力的支持.在这一背景下,搭载现代密码技术的各类硬件设备逐渐演变为不可或缺的生活基石.这些设备已经具备抵御传统密码分析的能力.近年来,学术界的研究重点之一是对设备在实际运行过程中产生的物理泄漏进行分析.这一领域被称为侧信道分析(Side-Channel Analysis,SCA).深度学习驱动的侧信道分析已被广泛认可为一种有效的方法,针对当前神经网络模型的功耗曲线数量需求大、鲁棒性差和收敛速度慢等问题,本文提出一种基于CNNbest的多尺度特征融合侧信道分析方法.首先,重构特征提取网络结构,以解决深层特征向量容易过度解释噪声细节的问题和模型过拟合问题.而后,使用滤波器阵列执行离散小波变换(Discrete Wavelet Transform,DWT)分析方法构造多解析度时频,提升数据质量.最后,引入轻量级的结合通道空间的卷积块注意力模块(Convolutional Block Attention Module,CBAM),以提高功耗曲线关键特征的学习效率.实验结果表明,本文方法对侧信道分析所需的功耗曲线较原模型减少了88.27%,显著提高了分析性能,能够满足侧信道建模和分析的要求.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0490-6756  
入藏号:  
CSCD:7745391  
Indexed Date:  
2024-08-30  
Institution Address:  
北京联合大学, 北京市信息服务工程重点实验室, 北京 100101, 中国  
Institution Address:  
中国科学院半导体研究所, 北京 100085, 中国  
Institution Address:  
北京联合大学机器人学院;;北京联合大学多智能体系统研究中心, ;;, ;;, 北京;;北京 100101;;100101, 中国  
Institution Address:  
北京微电子技术研究所, 北京 100076, 中国  
Institution Address:  
Beijing Union University, Beijing Key Laboratory of Information Service Engineering, Beijing 100101, China  
Institution Address:  
Institute of Semiconductors,Chinese Academy of Sciences, Beijing 100085, China  
Institution Address:  
College of Robotics,Beijing Union University;;Research Centre for Multi-Intelligent Systems,Beijing Union University, ;;, ;;, Beijing;;Beijing 100101;;100101  
Institution Address:  
Beijing Microelectronics Technology Institution, Beijing 100076, China  
Institution:  
北京联合大学  
Institution:  
中国科学院半导体研究所  
Institution:  
北京联合大学机器人学院;;北京联合大学多智能体系统研究中心  
Institution:  
北京微电子技术研究所  
Institution:  
Beijing Union University  
Institution:  
Institute of Semiconductors,Chinese Academy of Sciences  
Institution:  
College of Robotics,Beijing Union University;;Research Centre for Multi-Intelligent Systems,Beijing Union University  
Institution:  
Beijing Microelectronics Technology Institution

### 189. Efficient and Accurate Multi-Scale Topological Network for Single Image

摘要: Single image dehazing is a challenging ill-posed problem that has drawn significant attention in the last few years. Recently, convolutional neural networks have achieved great success in image dehazing. However, it is still difficult for these increasingly complex models to recover accurate details from the hazy image. In this paper, we pay attention to the feature extraction and utilization of the input image itself. To achieve this, we propose a Multi-scale Topological Network (MSTN) to fully explore the features at different scales. Meanwhile, we design a Multi-scale Feature Fusion Module (MFFM) and an Adaptive Feature Selection Module (AFSM) to achieve the selection and fusion of features at different scales, so as to achieve progressive image dehazing. This topological network provides a large number of search paths that enable the network to extract abundant image features as well as strong fault tolerance and robustness. In addition, ASFM and MFFM can adaptively select important features and ignore interference information when fusing different scale representations. Extensive experiments are conducted to demonstrate the superiority of our method compared with state-of-the-art methods.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
28  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
29  
ISSN:  
1520-9210  
eISSN:  
1941-0077  
入藏号:  
WOS:000809408000034  
Indexed Date:  
2022-06-20  
Institution Address:  
East China Normal Univ, Sch Comp Sci & Technol, Shanghai 200062, Peoples R China  
Institution Address:  
East China Normal Univ, Key Lab Adv Theory & Applicat Stat & Data Sci MOE, Shanghai 200062, Peoples R China  
Institution Address:  
Jiangxi Normal Univ, Sch Comp & Informat Engn, Nanchang 330022, Jiangxi, Peoples R China  
Institution:  
Institution:  
Institution:

### 190. Refine-FPN: Instance Segmentation Based on a Non-local Multi-feature

摘要: Rational use of multilevel structures of deep networks to extract multiscale features is crucial for instance segmentation. The Feature Pyramid Network (FPN) is a classical architecture that enriches the semantic information of multiscale objects. However, inherent defects in FPN structure are bound to cause loss of information during feature extraction and feature fusion. In this paper, we propose a feature pyramid structure (called Refine-FPN) based on a non-local multi-feature aggregation operation, a module that integrates multi-scale feature to rely on attention mechanisms to improve pyramid feature representation. The algorithm enriches the feature details of feature layers by aggregating multiple features to form a contextual global feature representation. By replacing FPN with Refine-FPN in the Mask R-CNN, our model improved the performance of the mask AP by 0.6% and 0.5% on the COCO dataset, when using ResNet-50 and ResNet-101 as the backbone, respectively. Moreover, it is friendly to integrate the proposed method into other popular architectures. For example, equipping the Cascade Mask R-CNN with Refine-FPN achieves an improvement of 0.5% and 0.4% mask AP under ResNet-50 and ResNet-101, respectively.  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1370-4621  
eISSN:  
1573-773X  
入藏号:  
WOS:000844947500001  
Indexed Date:  
2022-09-01  
Institution Address:  
Wuhan Univ Sci & Technol, Coll Informat Sci & Engn, Wuhan 430081, Peoples R China  
Institution:

### 191. EfficientSkinSegNet: a lightweight convolutional neural network for

摘要: Accurate segmentation of skin lesions is crucial for the early detection and treatment of skin cancer. In this study, we propose EfficientSkinSegNet, a novel lightweight convolutional neural network architecture specifically designed for precise skin lesion segmentation. EfficientSkinSegNet incorporates efficient feature extraction encoders and decoders, leveraging multi-head convolutional attention and spatial channel attention mechanisms to extract and enhance informative features while eliminating redundant ones. Furthermore, a multi-scale feature fusion module is introduced in the skip connections to facilitate effective fusion of features at different scales. Experimental evaluations on benchmark datasets demonstrate that EfficientSkinSegNet outperforms state-of-the-art methods in terms of segmentation accuracy while maintaining a compact model size. The proposed network shows promise for practical clinical diagnostic applications, providing a balance between segmentation performance and computational efficiency. Future research will focus on evaluating EfficientSkinSegNet's performance on diverse semantic segmentation tasks and optimizing it for medical image analysis.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0031-8949  
eISSN:  
1402-4896  
入藏号:  
WOS:001248922400001  
Indexed Date:  
2024-06-23  
Institution Address:  
Beijing Inst Petrochem Technol, Qingyuan North Rd 19, Beijing 102617, Peoples R China  
Institution Address:  
Air Force Med Univ, Air Force Med Ctr, Dept Radiol, 30 Fucheng Rd, Beijing 100142, Peoples R China  
Institution:  
Beijing Inst Petrochem Technol  
Institution:

### 192. A multi-scale feature fusion network-based fault diagnosis method for

摘要: A fault diagnosis method based on a multi-scale feature fusion network (MSFF-CNN) is proposed for the problem that the vibration signals of wind turbine bearings are easily disturbed by noise, and feature extraction is harrowing. Compared with the traditional diagnosis method, which has two stages of manual feature extraction and fault classification, this method combines the two into one. First, based on the characteristics of the bearing vibration signal, the multi-scale kernel algorithm is used to learn features in parallel at different scales. Then, the features extracted at different scales are fused to obtain complementary and rich diagnostic information. Finally, the Softmax classifier is used to output the fault diagnosis results. The simulation is carried out through the bearing vibration data of Case Western Reserve University. The results show that the accuracy of bearing fault diagnosis reaches 99.17%, proving the proposed method's high accuracy and effectiveness.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
0309-524X  
eISSN:  
2048-402X  
入藏号:  
WOS:000829117100001  
Indexed Date:  
2022-07-31  
Institution Address:  
North China Elect Power Univ, State Key Lab New Energy Power Syst, Baoding, Hebei, Peoples R China  
Institution:

### 193. Multi-scale convolution underwater image restoration network

摘要: Due to the complex underwater imaging environment and illumination conditions, underwater images have some quality degradation problems, such as low contrast, color distortion, texture blur and uneven illumination, which seriously restrict the application in underwater work. In order to solve these problems, we proposed a multi-scale feature fusion CNN based on underwater imaging model in this paper called Multi-Scale Convolution Underwater Image Restoration Network (MSCUIR-Net). Unlike most previous models that estimated the background light and transmittance, respectively, our model unifies the two parameters into one, predicts the univariate linear physical model through lightweight CNN, and directly generates end-to-end clean images. Based on the underwater imaging model, we synthesized the underwater image training set can simulate the shallow water to deep water environment. Then, we do experiments on synthetic images and real underwater images, and prove the superiority of this method through image evaluation indexes. The experimental results show that MSCUIR-Net has a good effect on underwater image restoration.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
9  
ISSN:  
0932-8092  
eISSN:  
1432-1769  
入藏号:  
WOS:000850577900001  
Indexed Date:  
2022-09-11  
Institution Address:  
Shanghai Univ, Sch Mechatron Engn & Automat, 99 Shangda Rd, Shanghai, Peoples R China  
Institution:

### 194. CNN-Based Feature Fusion Motor Fault Diagnosis

摘要: Artificial intelligence fields have been using deep learning in recent years. Due to its powerful data mining capabilities, deep learning has a wide-ranging impact on the diagnosis of motor faults. A method for diagnosing motor faults based on the multi-feature fusion of convolutional neural network (CNN) is presented in this paper. As far as the method is concerned, CNN is used as the basic framework, and the CNN model has been improved. First, the collected vibration and current signals are preprocessed. Second, segmented multi-time window synchronous input is performed on the processed data. In addition, a multi-scale feature extraction process and time series fusion of vibration and current signals subject to synchronous input in the same time window can be performed, which ultimately enables the identification of motor faults with a high degree of accuracy. In order to verify the validity of the proposed fault diagnosis model, an experimental platform for fault simulation was built for the motor, and vibration and current signals of different motor states were collected and verified by experimentation. According to the results of the experiment, the method can effectively combine motor vibration and current signal fault features, and thus motor fault diagnosis can be improved. In comparison with a single signal input, a multi-signal input provides greater accuracy and stability. As compared to other multi-signal feature fusion methods, such a deep learning model is able to extract fault features in a more comprehensive manner, which helps to improve the accuracy of motor fault diagnosis.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
11  
被引频次合计:  
11  
eISSN:  
2079-9292  
入藏号:  
WOS:000852541900001  
Indexed Date:  
2022-09-17  
Institution Address:  
Shanghai Dianji Univ, Sch Elect Engn, Shanghai 201306, Peoples R China  
Institution Address:  
Shanghai Dianji Univ, Sch Kaiserslautern Intelligent Mfg, Shanghai 201306, Peoples R China  
Institution:  
Institution:

### 195. Extraction of Agricultural Fields via DASFNet with Dual Attention

摘要: Agricultural fields are essential in providing human beings with paramount food and other materials. Quick and accurate identification of agricultural fields from the remote sensing images is a crucial task in digital and precision agriculture. Deep learning methods have the advantages of fast and accurate image segmentation, especially for extracting the agricultural fields from remote sensing images. This paper proposed a deep neural network with a dual attention mechanism and a multi-scale feature fusion (Dual Attention and Scale Fusion Network, DASFNet) to extract the cropland from a GaoFen-2 (GF-2) image of 2017 in Alar, south Xinjiang, China. First, we constructed an agricultural field segmentation dataset from the GF-2 image. Next, seven evaluation indices were selected to assess the extraction accuracy, including the location shift, to reveal the spatial relationship and facilitate a better evaluation. Finally, we proposed DASFNet incorporating three ameliorated and novel deep learning modules with the dual attention mechanism and multi-scale feature fusion methods. The comparison of these modules indicated their effects and advantages. Compared with different segmentation convolutional neural networks, DASFNet achieved the best testing accuracy in extracting fields with an F1-score of 0.9017, an intersection over a union of 0.8932, a Kappa coefficient of 0.8869, and a location shift of 1.1752 pixels. Agricultural fields can be extracted automatedly and accurately using DASFNet, which reduces the manual record of the agricultural field information and is conducive to further farmland surveys, protection, and management.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
19  
在 BIOSIS Citation Index 中的被引频次:  
3  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
21  
eISSN:  
2072-4292  
入藏号:  
WOS:000794625200001  
Indexed Date:  
2022-05-21  
Institution Address:  
Zhejiang Univ, Inst Agr Remote Sensing & Informat Technol Applic, Coll Environm & Resource Sci, Hangzhou 310058, Peoples R China  
Institution Address:  
Consolidat & Rehabil Ctr, Hangzhou 310007, Peoples R China  
Institution Address:  
Inst Cyber Syst & Control, Yuquan Campus, Hangzhou 310027, Zhejiang, Peoples R China  
Institution:  
Institution:  
Consolidat & Rehabil Ctr  
Institution:  
Inst Cyber Syst & Control

### 196. Source Camera Identification Algorithm Based on Multi-Scale Feature

摘要: The widespread availability of digital multimedia data has led to a new challenge in digital forensics. Traditional source camera identification algorithms usually rely on various traces in the capturing process. However, these traces have become increasingly difficult to extract due to wide availability of various image processing algorithms. Convolutional Neural Networks (CNN)-based algorithms have demonstrated good discriminative capabilities for different brands and even different models of camera devices. However, their performances is not ideal in case of distinguishing between individual devices of the same model, because cameras of the same model typically use the same optical lens, image sensor, and image processing algorithms, that result in minimal overall differences. In this paper, we propose a camera forensics algorithm based on multi-scale feature fusion to address these issues. The proposed algorithm extracts different local features from feature maps of different scales and then fuses them to obtain a comprehensive feature representation. This representation is then fed into a subsequent camera fingerprint classification network. Building upon the Swin-T network, we utilize Transformer Blocks and Graph Convolutional Network (GCN) modules to fuse multi-scale features from different stages of the backbone network. Furthermore, we conduct experiments on established datasets to demonstrate the feasibility and effectiveness of the proposed approach.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001315133800005  
Indexed Date:  
2024-09-23  
Institution Address:  
Hangzhou Dianzi Univ, Sch Comp Sci & Technol, Hangzhou 310018, Peoples R China  
Institution Address:  
Hangzhou Dianzi Univ, Shangyu Inst Sci & Engn, Shaoxing 312300, Peoples R China  
Institution Address:  
Zhejiang Police Coll, Minist Publ Secur, Key Lab Publ Secur Informat Applicat Based Big Dat, Hangzhou 310000, Peoples R China  
Institution Address:  
Menoufia Univ, Fac Artificial Intelligence, Shibin Al Kawm 32511, Egypt  
Institution:  
Institution:  
Institution:  
Institution:

### 197. EMFF-Net: Edge-Enhancement Multi-Scale Feature Fusion Network

摘要: According to medical research, colorectal polyps are considered typical precancerous lesions, making colonoscopic polyp images crucial for the early diagnosis of rectal cancer. However, variations in polyp size and shape, texture inconsistencies, and boundary ambiguities often pose significant challenges for polyp segmentation. To address these issues, we propose a multi-scale feature fusion network based on edge enhancement. Specifically, we utilize multi-scale feature fusion in each feature layer, where the extracted features are fused. From these extracted features, we generate a global mapping graph as a bootstrap region. Additionally, we introduce Spatial Channel Convolution (SCEConv) and Reverse Gated Channel Transformer (RGCT) to incorporate boundary information into the segmentation network. This approach enhances the layered features and produces a more refined segmentation map. Extensive qualitative and quantitative experiments on five benchmark datasets and two private datasets demonstrate that the EMFF-Net proposed in this paper significantly improves segmentation accuracy across six metrics. This represents a clear advantage over traditional CNNs and existing SOTA techniques, Especially, we achieved 81% mDice and 74% mIoU on the CVC-ColonDB dataset.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001420279300034  
Indexed Date:  
2025-02-23  
Institution Address:  
JiangNan Univ, Sch Artificial Intelligence & Comp Sci, Wuxi 214122, Jiangsu, Peoples R China  
Institution Address:  
Changshu Key Lab Med Artificial Intelligence & Big, Suzhou 215500, Jiangsu, Peoples R China  
Institution Address:  
Jiangnan Univ, Jiangsu Grad Workstat, Changshu 215500, Peoples R China  
Institution Address:  
Shanghai Univ, ULisboa Sch, Shanghai 200444, Peoples R China  
Institution Address:  
Soochow Univ, Changshu Affiliated Hosp, Dept Gastroenterol, Suzhou 215500, Jiangsu, Peoples R China  
Institution Address:  
Soochow Univ, Changshu Affiliated Hosp, Ctr Intelligent Med Technol Res, Suzhou 215500, Jiangsu, Peoples R China  
Institution:  
Institution:  
Changshu Key Lab Med Artificial Intelligence & Big  
Institution:  
Institution:  
Institution:  
Institution:

### 198. One-Stage Disease Detection Method for Maize Leaf Based on Multi-Scale

摘要: Plant diseases such as drought stress and pest diseases significantly impact crops' growth and yield levels. By detecting the surface characteristics of plant leaves, we can judge the growth state of plants and whether diseases occur. Traditional manual detection methods are limited by the professional knowledge and practical experience of operators. In recent years, a detection method based on deep learning has been applied to improve detection accuracy and reduce detection time. In this paper, we propose a disease detection method using a convolutional neural network (CNN) with multi-scale feature fusion for maize leaf disease detection. Based on the one-stage plant disease network YoLov5s, the coordinate attention (CA) attention module is added, along with a key feature weight to enhance the effective information of the feature map, and the spatial pyramid pooling (SSP) module is modified by data augmentation to reduce the loss of feature information. Three experiments are conducted under complex conditions such as overlapping occlusion, sparse distribution of detection targets, and similar textures and backgrounds of disease areas. The experimental results show that the average accuracy of the MFF-CNN is higher than that of currently used methods such as YoLov5s, Faster RCNN, CenterNet, and DETR, and the detection time is also reduced. The proposed method provides a feasible solution not only for the diagnosis of maize leaf diseases, but also for the detection of other plant diseases.  
在 BIOSIS Citation Index 中的被引频次:  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
20  
被引频次合计:  
20  
eISSN:  
2076-3417  
入藏号:  
WOS:000846330900001  
Indexed Date:  
2022-09-04  
Institution Address:  
Jimei Univ, Chengyi Univ Coll, Xiamen 361021, Peoples R China  
Institution Address:  
Jimei Univ, Sch Comp Engn, Xiamen 361021, Peoples R China  
Institution Address:  
Jimei Univ, Sch Ocean Informat Engn, Xiamen 361021, Peoples R China  
Institution:  
Institution:  
Institution:

### 199. A New Method Based on Deep Convolutional Neural Networks for Object

摘要: Accurate object detection and classification has a broad application in industrial tasks, such as fabric defect and invoice detection. Previous state-of-the-art methods such as SSD and Faster-RCNN usually need to carefully adjust anchor box related hyper parameters and have poor performance in special fields with large object size/ratio variations and complex background texture. In this study, we proposed a new accurate, robust, and anchor-free method to handle automatic object detection and classification problems. First, we used the feature pyramid network (FPN), to merge the feature maps of different scales of features extracted from a convolutional neural network (CNN), which allowed easy and robust multiscale feature fusion. Second, we built two subnets to generate candidate region proposals from the FPN outputs. followed by another CNN that determined the categories of the proposed regions from the two subnets.  
会议名称:  
2nd Artificial Intelligence on Fashion and Textile International Conference (AIFT) / 19th International Exhibition on Textile Industry (ShanghaiTex)  
会议日期:  
NOV 25-27, 2019  
会议地点:  
Shanghai, PEOPLES R CHINA  
会议赞助商:  
Hong Kong Polytechn Univ, Inst Textiles & Clothing  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2472-3444  
eISSN:  
2330-5517  
入藏号:  
WOS:000850266800006  
Indexed Date:  
2021-09-01  
Institution Address:  
Taikang Insurance Grp, Block B,Taikang Life Bldg,15b Fuxingmenrei St, Beijing, Peoples R China  
Institution:  
Taikang Insurance Grp

### 200. Remote Sensing Image Segmentation Network Based on Adaptive Multiscale

摘要: Remote sensing image segmentation algorithms are susceptible to interference from environmental factors,such as object occlusion and uneven illumination.Existing deep learning remote sensing image semantic segmentation methods usually adopt an end-to-end codec structure.However,they still suffer from inaccurate segmentation for the structure and contours of high similarity objects.Therefore,to improve the algorithm robustness and classification accuracy,a deep convolutional neural network remote sensing image semantic segmentation algorithm based on contour gradient learning is proposed.To improve the quality of the predicted feature maps,the adaptive attention-based multichannel multiscale feature fusion network (D-MMA Net) is proposed based on the SegNet model network.The DMA block uses an attention-based adaptive multiscale module to adaptively extract different scale features according to the learned weights to obtain more effective high level semantic features.To further refine the extracted object boundaries,the contour extraction module,a learnable contour extraction module,is proposed based on the principle of the Sobel edge detection operator.Finally,the contour information is combined with multi-scale semantic features to enhance the robustness of the spatial resolution of the image.The experimental results show that the proposed method improves the segmentation accuracy and produces good segmentation results for irregular object boundaries.  
摘要:  
遥感图像分割算法易受环境因素干扰,如物体遮挡、光照不均匀等。现有的深度学习遥感图像语义分割方法通常采取端到端的编解码结构,但针对相似度较高物体的结构和轮廓,仍存在分割不准确的问题。为了提高算法鲁棒性、分类准确率,提出一种基于轮廓梯度学习的深度卷积神经网络遥感图像语义分割算法。为了提高预测特征图的质量,首先基于SegNet模型,提出自适应注意力的多通道多尺度特征融合网络(D-MMA Net),其中D-MA block采用基于注意力的自适应多尺度模块,根据学习到的权重自适应地对不同尺度特征进行提取,以获得更多有效的高级语义特征。为进一步细化提取物体的边界,基于Sobel边缘检测算子原理提出可学习的轮廓提取模块。最后将轮廓信息与多尺度语义特征相结合,以增强对图像空间分辨率的鲁棒性。实验结果表明,所提算法提高分割的准确率,对于不规则物体边界,能有良好的分割效果。  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
CSCD:7404606  
Indexed Date:  
2023-06-02  
Institution Address:  
贵州大学计算机科学与技术学院, 贵阳, 贵州 550025, 中国  
Institution Address:  
贵阳欧比特宇航科技有限公司, 贵阳, 贵州 550027, 中国  
Institution Address:  
贵州师范学院大数据科学与智能工程研究院, 贵阳, 贵州 550018, 中国  
Institution Address:  
College of Computer Science and Technology,Guizhou University, Guiyang, Guizhou 550025, China  
Institution Address:  
Guiyang Orbita Aerospace Science&Technology Co.,Ltd., Guiyang, Guizhou 550027, China  
Institution Address:  
Big Data Science and Intelligent Engineering Research Institute,Guizhou Education University, Guiyang, Guizhou 550018, China  
Institution:  
贵州大学计算机科学与技术学院  
Institution:  
贵阳欧比特宇航科技有限公司  
Institution:  
贵州师范学院大数据科学与智能工程研究院  
Institution:  
College of Computer Science and Technology,Guizhou University  
Institution:  
Guiyang Orbita Aerospace Science&Technology Co.,Ltd.  
Institution:  
Big Data Science and Intelligent Engineering Research Institute,Guizhou Education University

### 201. Road vehicle detection and recognition algorithm based on densely

摘要: Regarding to the problems of low detection accuracy,poor real-time performance,and missed detection of small target vehicles in existing road vehicle detection and recognition algorithms,a road vehicle detection and recognition algorithm based on densely connected convolutional neural networks was proposed. Firstly,Based on YOLOv4(You Only Look Once version 4)network framework,by adopting the densely connected deep residual network structure,the feature reuse in the feature extraction stage was strengthened to realize the use of features with lower complexity on shallow layers. Then,a jump connection structure was integrated to the multi-scale feature fusion network to strengthen the feature information fusion and expression capability of the network,which reduced the missed detection rate of vehicles. Finally, the dimensional clustering algorithm was used to recalculate the anchor sizes,which were allocated to different detection scales according to a reasonable strategy. Experimental results show that the proposed algorithm achieves the detection accuracy of 98.21% and the detection speed of 48.05 frame/s on KITTI dataset,and it also has a good detection effect for vehicles in the complex and harsh environment of Berkeley DeepDrive(BDD100K)dataset,ensuring required real-time performance and effective accuracy improvement.  
摘要:  
针对现有道路车辆检测识别算法中存在的检测精度不高、实时性差以及小目标车辆漏检等问题,提出一种基于密集连接卷积神经网络的道路车辆检测与识别算法。首先,基于YOLOv4网络框架,通过采用密集连接的深度残差网络结构,加强特征提取阶段的特征复用,实现对浅层复杂度较低的特征的利用;然后,在多尺度特征融合网络引入跳跃连接结构,强化网络的特征信息融合和表征能力,以降低车辆漏检率;最后,采用维度聚类算法重新计算先验框尺寸,并按照合理的策略分配给不同检测尺度。实验结果表明,该算法在KITTI数据集上获得了98.21%的检测精度和48.05 frame/s的检测速度,对于BDD100K数据集中复杂恶劣环境中的车辆也有较好的检测效果,在满足实时检测要求的同时有效提升检测精度。  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1001-9081  
入藏号:  
CSCD:7181543  
Indexed Date:  
2022-06-10  
Institution Address:  
重庆交通大学交通运输学院, 重庆 400074, 中国  
Institution Address:  
School of Traffic and Transportation, Chongqing Jiaotong University, Chongqing 400074  
Institution:  
重庆交通大学交通运输学院  
Institution:  
School of Traffic and Transportation, Chongqing Jiaotong University

### 202. Attention-based multi-scale feature fusion for free-space detection

摘要: Free space detection is a very important task in road scene understanding. With the continued development of convolutional neural networks, free-space detection can be seen as a class-specific semantic segmentation problem. In this paper, a new encoding-decoding network structure-HRUnet is designed, which always maintains the input of high-resolution images in both the encoding and decoding phases. It extracts multi-scale information from RGB images and continuously fuses them, and finally achieves accurate spatial detection. In addition, in order to improve the accuracy of detection, the attention mechanism module-spin attention is proposed to achieve the interaction between channel and spatial dimensions when calculating channel attention, establish the come relationship between channel and space, reduce the loss of feature information, and further improve the accuracy of spatial detection. Experimental results show that the proposed neural network structure outperforms current popular models in terms of balanced the computational complexity and accuracy.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1751-956X  
eISSN:  
1751-9578  
入藏号:  
WOS:000793962600001  
Indexed Date:  
2022-05-20  
Institution Address:  
Shandong Technol & Business Univ, Coinnovat Ctr Shandong Coll & Univ Future Intelli, Sch Comp Sci & Technol, Yantai 264005, Peoples R China  
Institution Address:  
Aerosp New Generat Commun Co Ltd, Chongqing, Peoples R China  
Institution:  
Institution:  
Aerosp New Generat Commun Co Ltd

### 203. Image Super-Resolution Reconstruction Based on Attention Mechanism and

摘要: In super-resolution image reconstruction,most of the existing methods using Convolutional Neural Network (CNN) neglect the inherent attributes of natural images,and extract features only at a single scale. To address the problem,this paper proposes a network structure based on attention mechanism and multi-scale feature fusion.By using the attention mechanism,the non-local information and second-order features of the image are fused to improve the feature expression ability of the network.At the same time,different scales of convolutional kernels are used to extract different scales of information of the image,so as to preserve the complete information characteristics at different scales. Experimental results show that the reconstructed image by the proposed method outperforms Bicubic,SRCNN,SCN and LapSRN methods in terms of objective evaluation metrics and visual quality.  
摘要:  
目前多数利用卷积神经网络进行图像超分辨率重建的方法忽视对自然图像固有属性的捕捉,并且仅在单一尺度下提取特征。针对该问题,提出一种基于注意力机制和多尺度特征融合的网络结构。利用注意力机制融合图像的非局部信息和二阶特征,提高网络的特征表达能力,同时使用不同尺度的卷积核提取图像的不同尺度信息,以保存多尺度完整的信息特征。实验结果表明,该方法重建图像的客观评价指标和视觉效果均优于Bicubic、SRCNN、SCN和LapSRN方法。  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
5  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
1000-3428  
入藏号:  
CSCD:6917968  
Indexed Date:  
2021-05-28  
Institution Address:  
重庆邮电大学通信与信息工程学院, 重庆 400065, 中国  
Institution Address:  
School of Communication and Information Engineering,Chongqing University of Posts and Telecommunications, Chongqing 400065  
Institution:  
重庆邮电大学通信与信息工程学院  
Institution:  
School of Communication and Information Engineering,Chongqing University of Posts and Telecommunications

### 204. A Lightweight Vehicle-Pedestrian Detection Algorithm Based on Attention

摘要: Object detection is a critical technology of environmental perception for autonomous driving vehicle. The Convolutional Neural Network has gradually become a powerful tool in the field of vehicle detection because of its powerful ability of feature extraction. In aiming to reach the balance between speed and accuracy of detection in complex traffic scenarios, this paper proposes an improved lightweight and high-performance vehicle-pedestrian detection algorithm based on the YOLOv4. Firstly, the backbone network CSPDarknet53 is replaced by MobileNetv2 to reduce the number of parameters and raise the capability of feature extraction. Secondly, the method of multi-scale feature fusion is used to realize the information interaction among different feature layers. Finally, a coordinate attention mechanism is added to focus on the region of interest in the image by way of weight adjustment. The experimental results show that this improved model has a great performance in vehicle-pedestrian detection in traffic scenarios. Experimental results on PASCAL VOC datasets show that the improved model's mAP is 85.79% and speed is 35FPS, which has an increase of 4.31% and 16.7% compared to YOLOv4. Furthermore, the improved YOLOv4 model maintains a great balance between detection accuracy and speed on different datasets, indicating that it can be applied to vehicle-pedestrian detection in traffic scenarios.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
eISSN:  
1424-8220  
入藏号:  
WOS:000884067300001  
PubMed ID:  
36366178  
Indexed Date:  
2022-11-25  
Institution Address:  
Nanjing Forestry Univ, Sch Automot & Traff Engn, Nanjing 210037, Peoples R China  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Sch Energy & Power Engn, Nanjing 210016, Peoples R China  
Institution:  
Institution:

### 205. A Novel Faster RCNN Model Based on Multi-scale Feature Fusion and Shape

摘要: Compared with conventional targets, dense targets have less information, and the training data is difficult to be labeled, which leads to the poor detection effect with general target detection methods on dense targets, while the detection methods specially designed for dense targets are often too complex or not universal. To solve the above problems, this paper proposes a novel faster RCNN model based on multi-scale feature fusion and shape priori for dense vehicle detection. This proposed dense vehicle detection model is divided three steps. Firstly, the training strategy of multi-scale network captures the lost detail of target density. Secondly, the anchor point generation method based on shape priori is used to calculate the shape changes of dense targets at different scales. Thirdly, considering that the objects with dense distribution have large appearance differences, different sizes of convolution kernels are used to extract the feature information of different scales in images. It effectively solves the problem of target information loss in the existing detection models. Finally, we conduct experiments on the public CARPK dataset to illustrate the effectiveness of the proposed method. Compared with the state-of-the-art vehicle detection methods, the proposed network model can achieve better detection effect for the dense distributed targets in different scene conditions.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1557-2064  
eISSN:  
1557-2072  
入藏号:  
WOS:001003924600001  
Indexed Date:  
2023-06-20  
Institution Address:  
Zhengzhou Univ Sci & Technol, Sch Elect & Elect Engn, Zhengzhou 450064, Peoples R China  
Institution:  
Zhengzhou Univ Sci & Technol

### 206. Self-Supervised Real-World Image Denoising Based on Multi-Scale Feature

摘要: Deep learning denoising methods are often constrained by the high cost of acquiring real-world noisy images and the labor-intensive process of dataset construction. Our self-supervised Multi-Scale Blind-Spot Network with Adaptive Feature Fusion (MA-BSN) addresses these issues, offering an efficient solution for image denoising. MA-BSN mitigates the challenges of spatial noise correlation preservation and limited receptive fields, which are prevalent in existing self-supervised denoising approaches. The network employs a blind-spot architecture that generates sub-images at multiple scales, enhancing denoising beyond the capabilities of pixel-shuffle downsampling. A depth-wise convolutional Transformer network (DTN) extracts features across a global receptive field, addressing the convolutional neural networks' (CNNs) limitations. An adaptive feature fusion module (AFF) is introduced to refine feature learning for specific regions in the denoised images, leveraging attention mechanisms for improved performance. Our network's efficacy is validated through experiments on the SIDD and DND real-world noise benchmark datasets. Results on the DND dataset show a PSNR/SSIM of 38.41 dB/0.940, surpassing state-of-the-art self-supervised methods and underscoring our approach's superior denoising capability.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001199983200001  
Indexed Date:  
2024-04-15  
Institution Address:  
Qilu Normal Univ, Sch Informat Sci & Engn, Jinan 250200, Peoples R China  
Institution Address:  
Shandong Univ Engn & Vocat Technol, Sch Finance & Econ, Jinan 250200, Peoples R China  
Institution:  
Institution:  
Shandong Univ Engn & Vocat Technol

### 207. DBCGN: dual branch cascade graph network for skin lesion segmentation

摘要: Accurate segmentation of skin lesions in dermoscopic images is essential for early diagnosis and prevention of skin cancer. However, it is still a challenging task due to the large number of artifacts (hairs, bubbles, ink stains, etc.) in dermoscopic images and irregular shapes and blurred borders. In this paper, we propose a novel skin lesion segmentation network-DBCGN. In the feature extraction stage, We use the Transformer branch to build global contextual dependencies across scales against semantic features using three layers of high-level semantic features from the CNN branch. To more effectively couple local-global features, we design the Dual-branch Cascade-feature Fusion Module (DCFM) to implement multi-scale feature fusion of CNN and Transformer. In addition, to outline the boundaries of the target region more accurately, we design the Reverse Graph Reasoning Module (RGRM) which can refine the boundaries of the target region using low-level features. Extensive experiments have shown that our method outperforms the currently popular methods and provides more accurate segmentation of the target region. Furthermore, our method is computationally less complex and more efficient, which is more in line with the practical requirements of actual clinical segmentation.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1868-8071  
eISSN:  
1868-808X  
入藏号:  
WOS:000939388600001  
Indexed Date:  
2023-03-18  
Institution Address:  
Shandong Technol & Business Univ, Sch Comp Sci & Technol, Yantai, Peoples R China  
Institution Address:  
Shandong Univ Finance & Econ, Sch Comp Sci & Technol, Jinan, Peoples R China  
Institution:  
Institution:

### 208. A Multiscale Local-Global Feature Fusion Method for SAR Image

摘要: In recent years, the advancement of deep learning technology has led to excellent performance in synthetic aperture radar (SAR) automatic target recognition (ATR) technology. However, due to the interference of speckle noise, the task of classifying SAR images remains challenging. To address this issue, a multi-scale local-global feature fusion network (MFN) integrating a convolution neural network (CNN) and a transformer network was proposed in this study. The proposed network comprises three branches: a CovNeXt-SimAM branch, a Swin Transformer branch, and a multi-scale feature fusion branch. The CovNeXt-SimAM branch extracts local texture detail features of the SAR images at different scales. By incorporating the SimAM attention mechanism to the CNN block, the feature extraction capability of the model was enhanced from the perspective of spatial and channel attention. Additionally, the Swin Transformer branch was employed to extract SAR image global semantic information at different scales. Finally, the multi-scale feature fusion branch was used to fuse local features and global semantic information. Moreover, to overcome the problem of poor accuracy and inefficiency of the model due to empirically determined model hyperparameters, the Bayesian hyperparameter optimization algorithm was used to determine the optimal model hyperparameters. The model proposed in this study achieved average recognition accuracies of 99.26% and 94.27% for SAR vehicle targets under standard operating conditions (SOCs) and extended operating conditions (EOCs), respectively, on the MSTAR dataset. Compared with the baseline model, the recognition accuracy has been improved by 12.74% and 25.26%, respectively. The results demonstrated that Bayes-MFN reduces the inter-class distance of the SAR images, resulting in more compact classification features and less interference from speckle noise. Compared with other mainstream models, the Bayes-MFN model exhibited the best classification performance.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
3  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
4  
eISSN:  
2076-3417  
入藏号:  
WOS:001003767600001  
Indexed Date:  
2023-06-20  
Institution Address:  
Beijing Technol & Business Univ, Sch Artificial Intelligence, Beijing 100048, Peoples R China  
Institution Address:  
Beijing Technol & Business Univ, Key Lab Ind Internet & Big Data, China Natl Light Ind, Beijing 100048, Peoples R China  
Institution:  
Institution:

### 209. Rolling bearing fault diagnosis based on dilated convolution and

摘要: The traditional convolutional neural network (CNN) has the limitations of insufficiently extracting features from the original vibration signals and requiring a larger sensory field to fully capture the temporal correlation of the signals in the process of extracting the features when recognizing the fault types.A dilated convolution and enhanced multi-scale adaptive feature fusion model (DC-MAFFM) was proposed considering the inherent multi-scale characteristics of bearing vibration signals.The signal features were extracted using the large receptive field of the dilated convolution,and the residual connection was introduced to reduce the information loss on the convolution layer,so as to effectively filter the noise in the signal.An improved multi-scale feature extraction module was designed to capture complementary diagnostic features at different scales,meanwhile,the different-scale feature fusion was performed at each layer to fully learn the high-frequency and low-frequency features of the signal.The proposed feature adaptive fusion module was used to adaptively assign weights to the features at different scales to enhance the ability of discriminative feature learning.Verification was carried out on two bearing datasets,and results showed that the proposed model had strong diagnostic ability under noise and variable working conditions.In the case of strong noise,the fault diagnosis accuracy reached 88.08% and 75.56%,respectively,which demonstrated that the DC-MAFFM had a significant advantage over other methods.  
摘要:  
传统卷积神经网络(CNN)在识别故障类型时存在从原始振动信号中提取特征不足以及提取特征过程中需要更大的感受野以充分捕获信号的时间相关性的局限.针对轴承振动信号固有的多尺度特征,提出基于空洞卷积和增强型多尺度自适应特征融合的模型(DC-MAFFM).利用空洞卷积的大感受野提取信号特征,同时引入残差连接来减少卷积层上的信息损失,从而有效过滤信号中的噪声;设计改进的多尺度特征提取模块,在不同尺度上捕获互补的诊断特征,同时在各层都进行不同尺度特征融合,充分学习信号的高频和低频特征;利用提出的特征自适应融合模块对不同尺度的特征自适应赋予权重,增强判别特征学习的能力.在2个轴承数据集上进行验证,结果表明所提模型在噪声和变工况下有较强的诊断能力.在强噪声情况下,故障诊断准确率分别达到88.08%和75.56%,与其他方法相比有显著优势.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
1  
ISSN:  
1008-973X  
入藏号:  
CSCD:7729410  
Indexed Date:  
2024-08-08  
Institution Address:  
宁波大学机械工程与力学学院, 宁波, 浙江 315211, 中国  
Institution Address:  
Faculty of Mechanical Engineering and Mechanics,Ningbo University, Ningbo, Zhejiang 315211, China  
Institution:  
宁波大学机械工程与力学学院  
Institution:  
Faculty of Mechanical Engineering and Mechanics,Ningbo University

### 210. MMF-CNN: a multimodal fusion CNN network for winter wheat extraction

摘要: Timely and accurate acquisition of winter wheat planting areas is crucial for food security. In this study, Sentinel-1 and Sentinel-2 time-series data are integrated at the feature level to enhance the accuracy of winter wheat extraction. However, existing feature-level fusion models suffer from insufficient feature extraction and lack of feature completeness, thereby overlooking the complementarity and correlation between these two modalities. A Multimodal Fusion Convolutional Neural Network (MMF-CNN) model is proposed to address the issues above. Firstly, the images of Sentinel-1 and Sentinel-2 are processed to obtain the NDVI and backscatter characteristics of the winter wheat time series life cycle. A single feature and a combination of two features are then imported into each end of the model. The model adds a feature fusion module, which can fully extract the feature information. At the same time, the original features are retained in the process of multiscale feature fusion, which avoids the loss of the original information. Finally, Support Vector Machine (SVM), Long Short-Term Memory (LSTM), and Visual Geometry Group (VGG) are selected as the comparison models for comparative experiments, and the classification results of remote sensing images are obtained. These results demonstrate that the joint utilization of SAR and optical data yields the highest classification accuracy, with an F1 score of 97.42% for winter wheat. The overall accuracy (OA) of the proposed MMF-CNN method in this study is 96.87%, representing a 1.86% improvement compared to the Conv1D-CNN model. This improvement signifies adaptive feature learning at different hierarchical levels. Comparing the accuracy with other mainstream methods, the OA improves by 1.75%-4.37%, reveals finer ground details, and demonstrates faster performance. This study can provide methodological references for crop extraction studies based on multi-source data and time series analysis.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:001175826200001  
Indexed Date:  
2024-03-11  
Institution Address:  
China Univ Petr East China, Coll Oceanog & Space Informat, Qingdao, Peoples R China  
Institution Address:  
Land Surveying & Mapping Inst Shandong Prov, Satellite Applicat Ctr, Jinan, Peoples R China  
Institution Address:  
Land Surveying & Mapping Inst Shandong Prov, Satellite Applicat Ctr, 2301 Jingshi Rd, Jinan 250102, Peoples R China  
Institution:  
Institution:  
Land Surveying & Mapping Inst Shandong Prov  
Institution:  
Land Surveying & Mapping Inst Shandong Prov

### 211. Similar-Signal Recognition Method for φ-OTDR Systems Based on Multiscale

摘要: Objective A phase-sensitive optical time-domain reflectometer(phi-OTDR) system is a front monitoring and early warning technology that can acquire the location of disturbances in space and phase information of disturbances in time. With the advantages of high resolution, wide monitoring range, and strong anti-interference capability, this technology has been widely used in pipeline safety maintenance, intrusion warning, and large-equipment monitoring. However, due to the complex diversity of the application environment, the system suffers from low recognition accuracy and insufficient stability in actual use, particularly when similar signals are recognized in the system application. To solve these problems, this study proposes a similar-signal recognition method based on multiscale feature fusion. This method can effectively improve the recognition accuracy of similar signals while maintaining the recognition accuracy of the base signal. Methods The original signal is first decomposed into sub-signals in different frequency ranges using empirical mode decomposition (EMD) and wavelet packet decomposition (WPD). The original signal and individual sub-signals are then subjected to time-frequency feature extraction and approximate entropy feature extraction. The time-frequency features are used to evaluate the details of the time and frequency variations of the signal, the approximate entropy features are used to evaluate the complexity and regularity of the signal, and the multiscale signal decomposition and multi-feature extraction are used to amplify the feature differences between similar signals. Because the multiscale and multi-feature approach increases the dimensionality of the data, the proposed method utilizes principal component analysis (PCA) to combine high-dimensional features and reduce the dimensionality of system features, thereby improving system efficiency. Finally, the fused features are passed into a lightweight back-propagation (BP) neural network as input variables for signal data processing. Compared to other traditional neural networks, BP neural networks have the advantages of lightweight structures and high speed, enabling them to process signal data quickly. Results and Discussions Sub-signals decomposed by EMD and WPD have multiscale characteristics ranging from low to high frequencies. Each sub-signal contains a part of the signal domain within the main frequency-band range of the original data. Decomposition helps to amplify the feature gaps between different signals and facilitates subsequent multidimensional feature extraction (Fig.10). Following feature extraction and fusion, the four signals show significant differences in the feature space. Thus, even with a simple classifier, signal classification and recognition can be achieved (Fig.11). A comparison among extracting multi-features from original signal [Fig.12(a)], the CNN model [Fig.12(b)], and the multi-scale feature fusion[Fig.12(c)] reveals that the multi-scale feature fusion has higher recognition accuracy, where knocking and shaking-signal recognition accuracies reach 100% and trolleying and walking-signal recognition accuracies reach 98.5% and 98.0%,respectively. A comprehensive analysis reveals that the comprehensive recognition accuracy of the proposed method is increased by 8.4 and 9.0 percentage points over extracting multi-features from original signal and CNN model, respectively, and the similar-signal recognition accuracy is increased by 13.5 and 12.4 percentage points (Fig.13), respectively.  
These results verify that the method has high recognition accuracy. Conclusions Experimental results show that the decomposition method using EMD combined with WPD can obtain sub-signals at different scales. The time-frequency domain and approximate entropy features can in turn be extracted from the original signal and sub-signal to enhance the differentiation of similar-signal features more effectively. The PCA algorithm can then reduce the dimensionality of high-dimensional data, thus effectively reducing the number of training features. A well-designed six-layer lightweight BP neural network model can also effectively identify different types of signals when identifying signal features with significant differentiation. Compared with the extraction of features directly from the original signal, the proposed method can improve the integrated and similar-signal recognition accuracies by 8.4 and 13.5 percentage points, respectively. Compared to those of the CNN method, the overall recognition accuracy is improved by 9.0 percentage points, and the similar-signal recognition accuracy is improved by 14.3 percentage points. This method effectively improves similar-signal recognition while maintaining the recognition accuracy of underlying signals, which is of great value for expanding the applications of phi-OTDR systems  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
2  
ISSN:  
0258-7025  
入藏号:  
WOS:001190011100021  
Indexed Date:  
2024-04-17  
Institution Address:  
China Jiliang Univ, Coll Opt & Elect Technol, Hangzhou 310018, Zhejiang, Peoples R China  
Institution:

### 212. Crowd density estimation based on multi scale features fusion network

摘要: Deep learning has made substantial progress in crowd counting, but in practical applications, due to interference factors such as perspective distortion and complex background, the existing methods still have large errors in counting. In response to the above problems, this paper designs a multi-scale feature fusion network (IA-MFFCN) based on the reverse attention mechanism, which maps the image to the crowd density map for counting. The network consists of three parts: feature extraction module, inverse attention module, and back-end module. First, to overcome the problem of perspective distortion, deeper single-column CNNs was designed as a feature extraction module to extract multi-scale feature information and merge them; second, to avoid interference of complex backgrounds, the inverse attention module was designed, through the multi-scale inverse attention mechanism, reducing the influence of noise on counting accuracy. Finally, to generate a high-quality crowd density map, dilation convolution was introduced. Simultaneously, to enhance the sensitivity of the network to crowd counting, a comprehensive loss function based on Euclidean loss and predicted population loss is designed to improve training accuracy, to produce a more accurate density value. Experiments show that compared with the comparison algorithm, the algorithm in this paper has a significant reduction in the mean absolute error ( MAE) and mean square error (MSE) on the ShanghaiTech dataset, UCF\_CC\_50 dataset and WorldExpo'10 dataset.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
5  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000759316700001  
Indexed Date:  
2022-03-05  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao 266590, Peoples R China  
Institution Address:  
Shandong Univ Sci & Technol, Shandong Prov Key Lab Wisdom Mine Informat Techno, Qingdao 266590, Peoples R China  
Institution:  
Institution:

### 213. REDef-DETR: real-time and efficient DETR for industrial surface defect

摘要: Industrial surface defect detection is an important part of industrial production, which aims to identify and detecting various defects on the surface of product to ensure quality and meet customer requirements. With the development of deep learning and image processing technologies, the surface defect detection methods based on computer vision has become the mainstream method. However, the prevalent convolutional neural network-based defect detection methods also have many problems. For example, these methods rely on post-processing of Non-Maximum Suppression and have poor detection ability for small targets, which affects the speed and accuracy of surface defect detection in industrial scenarios. Therefore, we propose a novel DEtection TRansformer-based surface defect detection method. Firstly, we propose a Multi-scale Contextual Information Dilated module and fuse it into the backbone. The module is mainly composed of large kernel convolutions, which aims to expand the receptive field of the model, thus reducing the leakage rate of the model. Moreover, we design an efficient encoder which mainly contains two important modules, namely feature enhancement based on cascaded group attention module and efficient feature fusion module based on content-aware. The former module effectively enhances the high-level semantic information extracted by the backbone, thus enabling the model to better interpret features, and it can improve the problem of high computational cost of transformer encoder, thus increasing the detection speed. The latter module performs multi-scale feature fusion across the feature information of various scales, thus improving the detection accuracy of the model for small-size defects. Experimental results show that the proposed method achieves 80.6%mAP and 80.3FPS on NEU-DET, and 98.0%mAP and 79.4FPS on PCB-DET. Our proposed method exhibits excellent detection performance and achieves real-time and efficient surface defect detection capability to meet the needs of industrial surface defect detection.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001275391900001  
Indexed Date:  
2024-07-31  
Institution Address:  
Shenyang Univ Technol, Sch Informat Sci & Engn, 111 Shenliao West Rd, Shenyang 110870, Peoples R China  
Institution:

### 214. Building Extraction from High Spatial Resolution Imagery based on

摘要: Extraction of buildings from High Spatial Resolution Imagery (HSRI) plays an important role in remotely sensed imagery application. However, automatically extracting buildings from HSRI is still a challenging task due to factors such as large-scale variation of buildings, background complexity, and variation in appearance, etc. Especially, it is difficult in extracting both crowded small buildings and large buildings with accurate boundaries. To address these challenges, this paper presents an end-to-end encoder-decoder model to automatically extract buildings from HSRI. The designed network is called multiscale feature enhanced U-Shaped CNN with attention block and edge constraint (MAEU-CNN). Firstly, a Multiscale Feature Fusion (MFF) module is adopted in the encoder part of the network, which enables the network to aggregate features from multiple scales. Then, a Multi-scale Feature Enhancement module (MFEF) is added between the encoder and decoder parts to obtain multiscale receptive fields for obtaining multiscale context information. Thirdly, a dual attention mechanism is introduced to adaptively select representative feature maps for extraction of buildings instead of direct skipping connections. Lastly, in order to further solve the problem of segmentation result with poor boundaries aroused by the pooling operations in the MAEU-CNN, the geometric information of building boundary is introduced into the proposed MAEU-CNN by multi-task learning using the distance class map to produce fine-grained segmentations with precise boundaries. The performance of MAEU-CNN is examined through two different data sets at different building scales. The results show that MAEU-CNN obtains the greatest accuracy in each data set. The Precision, F1, and IoU is 93.4%, 93.62%, and 88.01%, respectively using the ISPRS Vaihingen semantic labeling contest data set. The Recall, F1, and IoU reach 95.45%, 95.58%, and 91.54%, respectively, using the WHU aerial image data set. Experimental results demonstrate that our proposed MAEU-CNN can achieve high accuracy for the extraction of building from remotely sensed imagery and show great robustness at different scales.  
摘要:  
从高空间分辨率图像(HSRI)中提取建筑物信息在遥感应用领域具有重要意义。然而,由于遥感影像中的建筑物尺度变化大、背景复杂和外观变化大等因素,从HSRI中自动提取建筑物仍然是一项具有挑战性的任务。特别是从影像中同时提取小型建筑物群和具有精确边界的大型建筑物时,难度更大。为解决这些问题,本文提出了一种端到端的编码器-解码器神经网络模型,用于从HSRI中自动提取建筑物。所设计的网络称为MAEU-CNN(Multiscale Feature Enhanced U-shaped CNN with Attention Block and Edge Constraint)。首先,在设计的网络编码部分加入多尺度特征融合(MFF)模块,使网络能够更好地聚集多个尺度特征。然后,在编码器和解码器部分之间添加了多尺度特征增强模块(MFEF),以获得不同尺寸的感受野,用于获取更多的多尺度上下文信息。在跳跃连接部分引入双重注意机制,自适应地选择具有代表性的特征图用于提取建筑物。最后,为了进一步解决MAEU-CNN中由于池化及卷积操作导致的分割结果边界模糊的问题,引入多任务学习机制,将建筑物的边界几何信息融入网络中以优化提取的建筑物边界,最终获得精确边界的建筑物信息。MAEU-CNN在ISPRS Vaihingen语义标记数据集和WHU航空影像数据集2种不同尺度建筑物数据集上进行了试验分析,在ISPRS Vaihingen语义标记数据集上,MAEU-CNN在精度、F1分数和IoU指标中获得了最高精度,分别达到了93.4%、93.62%和88.01%;在WHU航空影像数据集上,召回率、F1分数和IoU指标中也获得了最高精度,分别达到了95.45%、95.58%和91.54%。结果表明,本文所提出的MAEU-CNN从遥感图像中提取建筑物信息精度较高,并且对于不同尺度具有较强的鲁棒性。  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
1560-8999  
入藏号:  
CSCD:7228939  
Indexed Date:  
2022-09-03  
Institution Address:  
(徐州)中国矿业大学环境与测绘学院, 徐州, 江苏 221161, 中国  
Institution Address:  
香港理工大学土地测量及地理资讯学系, 999077, 香港  
Institution Address:  
School of Environment and Spatial Informatics, China University of Mining and Technology (Xuzhou), Xuzhou, Jiangsu 221161, China  
Institution Address:  
Department of Land Surveying and Geo-informatics, The Hong Kong Polytechnic University, Hongkong 999077, China  
Institution:  
(徐州)中国矿业大学环境与测绘学院  
Institution:  
香港理工大学土地测量及地理资讯学系  
Institution:  
School of Environment and Spatial Informatics, China University of Mining and Technology (Xuzhou)  
Institution:  
Department of Land Surveying and Geo-informatics, The Hong Kong Polytechnic University

### 215. FDA-FFNet: A Feature-Distance Attention-Based Change Detection Network

摘要: Convolutional neural networks have demonstrated remarkable capability in extracting deep semantic features from images, leading to significant advancements in various image processing tasks. This success has also opened up new possibilities for change detection (CD) in remote sensing applications. But unlike the conventional image recognition tasks, the performance of AI models in CD heavily relies on the method used to fuse the features from two different phases of the image. The existing deep-learning-based methods for CD typically fuse features of bitemporal images using difference or concatenation techniques. However, these approaches often fail tails to prioritize potential change areas adequately and neglect the rich contextual information essential for discerning subtle changes, potentially leading to slower convergence speed and reduced accuracy. To tackle this challenge, we propose a novel feature fusion approach called feature-difference attention-based feature fusion CD network. This method aims to enhance feature fusion by incorporating a feature-difference attention-based feature fusion module, enabling a more focused analysis of change areas. Additionally, a deep-supervised attention module is implemented to leverage the deep surveillance module for cascading refinement of change areas. Furthermore, an atrous spatial pyramid pooling fast is employed to efficiently acquire multiscale object information. The proposed method is evaluated on two publicly available datasets, namely the WHU-CD and LEVIR-CD datasets. Compared with the state-of-the-art CD methods, the proposed method outperforms in all metrics, with an intersection over union of 92.49% and 85.56%, respectively.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001140808700022  
Indexed Date:  
2024-02-04  
Institution Address:  
Wuhan Univ, Sch Remote Sensing & Informat Engn, Wuhan 430079, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Smart Cities Res Inst, Hong Kong, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Dept Land Surveying & Geoinformat, Hong Kong, Peoples R China  
Institution Address:  
China Univ Min & Technol, Sch Environm & Spatial Informat, Xuzhou 221116, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 216. BO-densenet: A bilinear one-dimensional densenet network based on

摘要: With the rapid development of deep learning techniques, convolutional neural networks have been widely used in the field of spectroscopy. In this paper, a bilinear branching Densenet network model (BO-Densenet) based on multi-scale feature fusion is constructed by applying a one-dimensional convolutional neural network to classify six woods: Tung wood, Balsa wood, Poplar wood, PVA-modified Poplar wood, Nano-silica-sol modified Poplar wood, and PVA-Nano-silica-sol modified Poplar wood. The results show that BO-Densenet achieves 98.90% accuracy in classification on the test set, which is higher than 82.09% of Partial Least Squares, and also higher than 89.88% of Lenet, 93.56% of Alexnet, 94.12% of Resnet-18 and 96.69% of Densenet-40 when compared with other deep learning algorithms. This shows that the BO-Densenet proposed in this paper can accurately achieve wood classification and has potential application prospects.  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0169-7439  
eISSN:  
1873-3239  
入藏号:  
WOS:001053525800001  
Indexed Date:  
2023-09-02  
Institution Address:  
Northeast Forestry Univ, Coll Sci, Harbin 150040, Peoples R China  
Institution:

### 217. DCNNet: A Distributed Convolutional Neural Network for Remote Sensing

摘要: With the development of information technology, multiplatform collaborative collection and processing of remote sensing (RS) images has become a significant trend. However, the existing models are challenging to achieve accurate and efficient image interpretation on RS multiplatform systems. To solve this problem, we propose a novel distributed convolutional neural network (DCNNet) and demonstrate the superiority of our method in RS image classification. First, a progressive inference mechanism is introduced to support most images to be classified in advance with satisfactory accuracy, which minimizes redundant cloud transmission and achieves higher inference acceleration. Meanwhile, a distributed self-distillation paradigm is designed to integrate and refine in-depth features, performing efficient knowledge transfer between the terminals and the cloud network. Second, a multiscale feature fusion (MSFF) module is presented to extract valid receptive fields and assign weights to crucial channel dimension features. Finally, a sampling augmentation (SA) attention is proposed to enhance the effective feature representation of RS images through a bottom-up and top-down feedforward structure. We conducted extensive experiments and visual analyses on three benchmark scene classification datasets and one fine-grained dataset. Compared with the existing methods, DCNNet consolidates several advantages in terms of accuracy, computation, transmission, and processing efficiency into a single framework for multiplatform RS image classification.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000966891200001  
Indexed Date:  
2023-05-22  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Aerosp Informat Res Inst, Chinese Acad Sci, Key Lab Network Informat Syst Technol NIST, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 100190, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 218. RMT-YOLOv9s: An Infrared Small Target Detection Method Based on UAV

摘要: Unmanned aerial vehicles (UAVs) and infrared imaging technology have numerous applications in civilian fields. To address the issues of low accuracy resulting from complex ground backgrounds, small target size, and limited target features in UAV remote sensing infrared image target detection, we use the YOLOv9s model and the latest retentive networks meet vision transformers (RMTs) technology and propose the RMT-YOLOv9s model for infrared small target detection. First, a convolutional neural network (CNN)-RMT-based backbone is proposed by incorporating the RMT model into the backbone network of YOLOv9s, which extracts both local and global features for small target detection. Then, an improved neck multiscale feature-fusion network RMTELAN-PANet is designed using the novel convolutional RMTELAN module proposed in this letter, which can better capture and use semantic information from feature maps. Finally, efficient multiscale attention (EMA) attention module and upsampling Dysample module are integrated into RMTELAN-PANet to further improve the feature information of small targets. Experiments on the HIT-UAV dataset show that RMT-YOLOv9s outperforms other popular methods in infrared small target detection.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001349449300003  
Indexed Date:  
2024-11-15  
Institution Address:  
Beijing Inst Technol, Sch Elect & Mech, Beijing 100081, Peoples R China  
Institution Address:  
Sci & Technol Electromech Dynam Control Lab, Xian 710065, Peoples R China  
Institution Address:  
Beijing Univ Technol, Sch Informat Sci & Technol, Beijing 100124, Peoples R China  
Institution:  
Institution:  
Sci & Technol Electromech Dynam Control Lab  
Institution:

### 219. Dual-feature Fusion Attention Network for Small Object Segmentation

摘要: Accurate segmentation of medical images is an important step during radiotherapy planning and clinical diagnosis. However, manually marking organ or lesion boundaries is tedious, time-consuming, and prone to error due to subjective variability of radiologist. Automatic segmentation remains a challenging task owing to the variation (in shape and size) across subjects. Moreover, existing convolutional neural networks based methods perform poorly in small medical objects segmentation due to class imbalance and boundary ambiguity. In this paper, we propose a dual feature fusion attention network (DFF-Net) to improve the segmentation accuracy of small objects. It mainly includes two core modules: the dual-branch feature fusion module (DFFM) and the reverse attention context module (RACM). We first extract multi-resolution features by multi-scale feature extractor, then construct DFFM to aggregate the global and local contextual information to achieve information complementarity among features, which provides sufficient guidance for accurate small objects segmentation. Moreover, to alleviate the degradation of segmentation accuracy caused by blurred medical image boundaries, we propose RACM to enhance the edge texture of features. Experimental results on datasets NPC, ACDC, and Polyp demonstrate that our proposed method has fewer parameters, faster inference, and lower model complexity, and achieves better accuracy than more state-of-the-art methods.  
Web of Science 核心合集中的 "被引频次":  
10  
在 BIOSIS Citation Index 中的被引频次:  
4  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:001006929700001  
PubMed ID:  
37178604  
Indexed Date:  
2023-06-26  
Institution Address:  
Chengdu Univ Informat Technol, Coll Comp Sci, Chengdu 610000, Peoples R China  
Institution Address:  
Xihua Univ, Sch Comp & Software Engn, Chengdu 610000, Peoples R China  
Institution Address:  
Univ Agr Faisalabad Pakistan, Agr Univ Rd, Faisalabad 38000, Pakistan  
Institution Address:  
Sichuan Univ, West China Hosp, Dept Crit Care Unit, Chengdu 610000, Peoples R China  
Institution Address:  
Chengdu Univ TCM, Chengdu Peoples Hosp 1, Dept Radiol, Integrated TCM & Western Med Hosp, Chengdu 610000, Peoples R China  
Institution Address:  
Sichuan Univ, West China Hosp, Chengdu 610000, Peoples R China  
Institution Address:  
Sichuan Univ, Coll Comp Sci, Chengdu 610000, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 220. Multilabel Remote Sensing Image Annotation With Multiscale Attention and

摘要: Deep-learning-based multilabel image annotation is receiving increasing attention in the field of remote sensing due to the great success of deep networks in single-label remote sensing image classification. Compared with those low-level features, the features extracted by the convolutional neural network (CNN) are more informative and can alleviate the problem of semantic gap. However, the CNN model tends to ignore the smaller objects when objects of different sizes exist in an image. In addition, how to efficiently leverage the correlation among multiple labels to enhance annotation performance remains an open issue. In this article, we propose an end-to-end deep learning framework for multilabel remote sensing image annotation. The framework is composed of a multiscale feature fusion module, a channel-spatial attention learning module, and a label correlation extraction module. The multiscale features from different layers of a CNN model are first fused and refined by using a channel-spatial attention mechanism. Then, the label correlation information is extracted from a label co-occurrence matrix and embedded into the multiscale attentive features to increase the discriminative ability of the resulting image features. The experiments on two benchmark datasets demonstrate the superiority of the proposed method in comparison with the state-of-the-art methods.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
23  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
23  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000678338200004  
Indexed Date:  
2021-08-05  
Institution Address:  
Shanghai Univ, Sch Commun & Informat Engn, Shanghai 200444, Peoples R China  
Institution:

### 221. An Efficient and Accurate 3D Multiple-Contextual Semantic Segmentation

摘要: Convolutional neural networks have become popular in medical image segmentation, and one of their most notable achievements is their ability to learn discriminative features using large labeled datasets. Two-dimensional (2D) networks are accustomed to extracting multiscale features with deep convolutional neural network extractors, i.e., ResNet-101. However, 2D networks are inefficient in extracting spatial features from volumetric images. Although most of the 2D segmentation networks can be extended to three-dimensional (3D) networks, extended 3D methods are resource and time intensive. In this paper, we propose an efficient and accurate network for fully automatic 3D segmentation. We designed a 3D multiple-contextual extractor (MCE) to simulate multiscale feature extraction and feature fusion to capture rich global contextual dependencies from different feature levels. We also designed a light 3D ResU-Net for efficient volumetric image segmentation. The proposed multiple-contextual extractor and light 3D ResU-Net constituted a complete segmentation network. By feeding the multiple-contextual features to the light 3D ResU-Net, we realized 3D medical image segmentation with high efficiency and accuracy. To validate the 3D segmentation performance of our proposed method, we evaluated the proposed network in the context of semantic segmentation on a private spleen dataset and public liver dataset. The spleen dataset contains 50 patients' CT scans, and the liver dataset contains 131 patients' CT scans.  
会议名称:  
43rd Annual International Conference of the IEEE-Engineering-in-Medicine-and-Biology-Society (IEEE EMBC)  
会议日期:  
NOV 01-05, 2021  
会议地点:  
ELECTR NETWORK  
会议赞助商:  
IEEE Engn Med & Biol Soc; IEEE; Elsevier; Inst Engn & Technol  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1557-170X  
eISSN:  
1558-4615  
ISBN:  
978-1-7281-1179-7  
入藏号:  
WOS:000760910503058  
PubMed ID:  
34891948  
Indexed Date:  
2022-04-21  
Institution Address:  
Ritsumeikan Univ, Grad Scool Informat Sci & Engn, Kusatsu, Shiga, Japan  
Institution Address:  
Yamaguchi Univ, Fac Sci, Yamaguchi, Japan  
Institution Address:  
Tokyo Metropolitan Univ, Tokyo, Japan  
Institution Address:  
Koseikai Takeda Hosp, Kyoto, Japan  
Institution:  
Institution:  
Institution:  
Institution:  
Koseikai Takeda Hosp

### 222. Face2Nodes: Learning facial expression representations with

摘要: Deep convolutional neural networks (CNNs) have become the standard model architecture for facial expression recognition (FER). However, CNN-based models struggle to capture the structural correlations between different local regions in a face image. Recent methods based on Vision Transformer (ViT) have been introduced to capture long-range dependencies among local regions. Nonetheless, ViT-based approaches are vulnerable to facial regions unrelated to expressions and may learn redundant correlation representations due to their self-attention mechanism. To address these issues, we propose a novel graph-based model called Face2Nodes, which can flexibly learn the graph representations of facial expressions without requiring additional auxiliary facial information such as landmarks. Our Face2Nodes consists of two key components: a multi-scale feature fusion-based patch embedding and a relation-aware dynamic graph convolution network. The patch embedding method uses a multi-scale feature fusion mechanism to obtain more discriminative graph node features for further graph representation learning. A dynamic graph is constructed using the dilated k-nearest neighbors algorithm, and a relation-aware graph convolution operator is designed to learn the latent informative correlations among different nodes in the graph. Extensive experiment results show that Face2Nodes achieves state-of-the-art performance on several popular in-the-wild FER datasets, with overall accuracies of 91.41%, 91.02%, and 66.69% on the FERPlus, RAF-DB, and AffectNet databases, respectively. Furthermore, we found that CNN-based FER approaches have a more significant performance gap between pre-training and training from scratch than Face2Nodes, demonstrating that our model is more data-efficient than CNN-based approaches.  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
7  
ISSN:  
0020-0255  
eISSN:  
1872-6291  
入藏号:  
WOS:001071752200001  
Indexed Date:  
2023-10-12  
Institution Address:  
Key Lab Intelligent Educ Technol & Applicat Zhejia, Jinhua 321004, Zhejiang, Peoples R China  
Institution Address:  
Jinan Univ, Coll Informat Sci & Technol, Dept Comp Sci, Guangzhou 510632, Peoples R China  
Institution Address:  
Jinan Univ, Guangdong Inst Smart Educ, Guangzhou 510632, Peoples R China  
Institution:  
Key Lab Intelligent Educ Technol & Applicat Zhejia  
Institution:  
Institution:

### 223. LRTransDet: A Real-Time SAR Ship-Detection Network with Lightweight ViT

摘要: In recent years, significant strides have been made in the field of synthetic aperture radar (SAR) ship detection through the application of deep learning techniques. These advanced methods have substantially improved the accuracy of ship detection. Nonetheless, SAR images present distinct challenges, including complex backgrounds, small ship targets, and noise interference, thereby rendering the detectors particularly demanding. In this paper, we introduce LRTransDet, a real-time SAR ship detector. LRTransDet leverages a lightweight vision transformer (ViT) and a multi-scale feature fusion neck to address these challenges effectively. First, our model implements a lightweight backbone that combines convolutional neural networks (CNNs) and transformers, thus enabling it to simultaneously capture both local and global features from input SAR images. Moreover, we boost the model's efficiency by incorporating the faster weighted feature fusion (Faster-WF2) module and coordinate attention (CA) mechanism within the feature fusion neck. These components optimize computational resources while maintaining the model's performance. To overcome the challenge of detecting small ship targets in SAR images, we refine the original loss function and use the normalized Wasserstein distance (NWD) metric and the intersection over union (IoU) scheme. This combination improves the detector's ability to efficiently detect small targets. To prove the performance of our proposed model, we conducted experiments on four challenging datasets (the SSDD, the SAR-Ship Dataset, the HRSID, and the LS-SSDD-v1.0). The results demonstrate that our model surpasses both general object detectors and state-of-the-art SAR ship detectors in terms of detection accuracy (97.8% on the SSDD and 93.9% on the HRSID) and speed (74.6 FPS on the SSDD and 75.8 FPS on the HRSID), all while demanding 3.07 M parameters. Additionally, we conducted a series of ablation experiments to illustrate the impact of the EfficientViT, the Faster-WF2 module, the CA mechanism, and the NWD metric on multi-scale feature fusion and detection performance.  
Web of Science 核心合集中的 "被引频次":  
9  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
9  
eISSN:  
2072-4292  
入藏号:  
WOS:001120533500001  
Indexed Date:  
2024-01-09  
Institution Address:  
Peking Univ, Sch Software & Microeletron, Beijing 102600, Peoples R China  
Institution Address:  
Peking Univ, Sch Integrated Circuits, Beijing 100871, Peoples R China  
Institution Address:  
Beijing Aerosp Automat Control Inst, Beijing 100039, Peoples R China  
Institution:  
Institution:  
Institution:  
Beijing Aerosp Automat Control Inst

### 224. A feature enhancement FCOS algorithm for dynamic traffic object

摘要: The development of object detection plays an important role in the realisation of fully autonomous driving, and the feature extraction is the key step for object detection. There has been significant difference and scale variation of object features for different road traffic participants (RTPs), meanwhile traditional Convolutional Neural Networks (CNNs) was difficult to extract object features efficiently for small targets. In order to improve the ability of feature extraction, a RTP object detection method combining dynamic convolution and feature enhancement was proposed. The Fully Convolutional One-Stage (FCOS) object detection algorithm was used as baseline. First, the dynamic convolution module was designed in the backbone network to identify different object features to the maximum extent. Second, a dual attention module was designed to filter object feature information while reducing the amount of computation. Finally, in the detection part, the feature expression ability of shallow network was further enhanced by multi-scale feature fusion module, and the effectiveness of the proposed algorithm was verified using Cityscapes dataset. The experimental result indicated that mAP increased by 2.3% compared with baseline. This study can improve the efficiency of RTP detection and contribute to the industrialisation of intelligent connected vehicles.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0954-0091  
eISSN:  
1360-0494  
入藏号:  
WOS:001176322400001  
Indexed Date:  
2024-03-22  
Institution Address:  
East China Jiaotong Univ, Sch Transportat Engn, Nanchang, Peoples R China  
Institution Address:  
Tsinghua Univ, Suzhou Automot Res Inst, Suzhou, Peoples R China  
Institution Address:  
Tsinghua Univ, Suzhou Automot Res Inst, Suzhou 215200, Peoples R China  
Institution:  
Institution:  
Institution:

### 225. Rolling bearing fault diagnosis method based on MTF-MFACNN

摘要: A rolling bearing fault diagnosis method based on the Markov transition field (MTF) and multi-scale feature aggregation convolutional neural network (MFACNN) is proposed to address the problems of excessive parameter number, slow training speed, and insufficient generalization of traditional CNNs. Firstly, the original vibration signal is input into the MTF and converted into two-dimensional images with time correlation. Then, in order to effectively aggregate feature information at different scales and levels, a MFA module is presented to capture rich information from feature maps at different scales and assign different weights to these features for fusion. Secondly, while ensuring the lightweight of the model, utilizing feature information of different resolutions, a lightweight feature fusion module is put forward to fuse multiple feature maps together to improve the performance and efficiency of the model. On this basis, an MFACNN model is constructed. Finally, the two-dimensional images are input into MTF-MFACNN and experimentally validated using two different datasets. The results show that the proposed method has faster calculation speed, higher fault recognition accuracy, and stronger generalization performance compared to other methods.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001116567600001  
Indexed Date:  
2023-12-16  
Institution Address:  
Lanzhou Univ Technol, Sch Mech & Elect Engn, Lanzhou 730050, Gansu, Peoples R China  
Institution:

### 226. MSSIF-Net: an efficient CNN automatic detection method for freight train

摘要: Freight trains are one of the most important modes of transportation. The fault detection of freight train parts is crucial to ensure the safety of train operation. Given the low detection efficiency and accuracy of traditional train fault detection methods, a novel one-stage object detection method called the multi-scale spatial information fusion CNN network (MSSIF-Net) based on YOLOv4 is proposed in this study. The adaptive spatial feature fusion method and multi-scale channel attention mechanism are used to construct the multi-scale feature sharing network and consequently realize feature information sharing at different levels and promote detection accuracy. The mean average precision values of MSSIF-Net on the train image test set, PASCAL VOC 2007 test set, and surface defect detection dataset are 94.73%, 87.76%, and 75.54%, respectively, outperforming YOLOv4, Faster R-CNN, CenterNet, RetinaNet, and YOLOX-l. The detection speed of MSSIF-Net is 33.10 FPS, achieving a good balance between detection accuracy and speed. In addition, the MSSIF-Net performance is estimated after adding noise or rotating the train images at a slight angle to simulate a real scene. Experimental results indicate that MSSIF-Net has favorable anti-interference ability.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
ISSN:  
0941-0643  
eISSN:  
1433-3058  
入藏号:  
WOS:000890131800003  
Indexed Date:  
2022-12-14  
Institution Address:  
Hunan Univ Technol, Coll Comp Sci, Zhuzhou 412007, Hunan, Peoples R China  
Institution Address:  
Hunan Univ Technol & Business, Sch Comp Sci, Changsha 410205, Hunan, Peoples R China  
Institution Address:  
SUNY Coll New Paltz, Dept Comp Sci, 100 Hawk Dr, New Paltz, NY 12561 USA  
Institution:  
Institution:  
Institution:

### 227. Scale-aware attention-based multi-resolution representation for

摘要: The performance of multi-person pose estimation has significantly improved with the development of deep convolutional neural networks. However, two challenging issues are still ignored but are key factors causing deterioration in the keypoint localization. These two issues are scale variation of human body parts and huge information loss caused by consecutive striding in multiple upsampling. In this paper, we present a novel network named 'Scale-aware attention-based multi-resolution representation network' (SaMr-Net) which targets to make the proposed method against scale variation and prevent the detail information loss in upsampling, leading more precisely keypoint estimation. The proposed architecture adopts the high-resolution network (HRNet) as the backbone, we first introduce dilated convolution into the backbone to expand the receptive field. Then, attention-based multi-scale feature fusion module is devised to modify the exchange units in the HRNet, allowing the network to learn the weights of each fusion component. Finally, we design a scale-aware keypoint regressor model that gradually integrates features from low to high resolution, enhancing the invariance in different scales of pose parts keypoint estimation. We demonstrate the superiority of the proposed algorithm over two benchmark datasets: (1) the MS COCO keypoint benchmark, and (2) the MPII human pose dataset. The comparison shows that our approach achieves superior results.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
0942-4962  
eISSN:  
1432-1882  
入藏号:  
WOS:000656926300001  
Indexed Date:  
2021-06-12  
Institution Address:  
Shaanxi Normal Univ, Key Lab Modern Teaching Technol, Minist Educ, Xian 710062, Peoples R China  
Institution Address:  
Shaanxi Normal Univ, Sch Comp Sci, Xian 710062, Peoples R China  
Institution:  
Institution:

### 228. Multiscale attention-based detection of tiny targets in aerial beach

摘要: Tiny target detection in marine scenes is of practical importance in marine vision applications such as personnel search and rescue, navigation safety, and marine management. In the past few years, methods based on deep convolutional neural networks (CNN) have performed well for targets of common sizes. However, the accurate detection of tiny targets in marine scene images is affected by three difficulties: perspective multiscale, tiny target pixel ratios, and complex backgrounds. We proposed the feature pyramid network model based on multiscale attention to address the problem of tiny target detection in aerial beach images with large field-of-view, which forms the basis for the tiny target recognition and counting. To improve the ability of the tiny targets' feature extraction, the proposed model focuses on different scales of the images to the target regions based on the multiscale attention enhancement module. To improve the effectiveness of tiny targets' feature fusion, the pyramid structure is guided by the feature fusion module in order to give further semantic information to the low-level feature maps and prevent the tiny targets from being overwhelmed by the information at the high-level. Experimental results show that the proposed model generally outperforms existing models, improves accuracy by 8.56 percent compared to the baseline model, and achieves significant performance gains on the TinyPerson dataset. The code is publicly available via Github.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2296-7745  
入藏号:  
WOS:000901572000001  
Indexed Date:  
2023-01-10  
Institution Address:  
Beijing Informat Sci & Technol Univ, Inst Appl Math, Beijing, Peoples R China  
Institution Address:  
Beijing Informat Sci & Technol Univ, Key Lab Minist Educ Optoelect Measurement Technol, Beijing, Peoples R China  
Institution:  
Institution:

### 229. Side-Scan Sonar Image Segmentation Based on Multi-Channel Fusion

摘要: Side-scan sonar is an important application in the field of ocean exploration. Accurate segmentation of target regions in side-scan sonar images is a challenging issue due to the low-resolution and strong noise interference. To accurately and faster segment the different categories target in sonar image, a novel convolutional neural networks (CNNs) model is proposed in this study. Firstly, the deep separable residual module is used for target regions multi-scale feature extraction and suppression noise feature information interference, and the multi-channel feature fusion method is used to enhance feature information transfer of convolution layers. Secondly, the adaptive supervised function is used for pixel-wise classification of different categories targets. Finally, to improve model generalization ability and robustness, the adaptive transfer learning method is introduced in the model training process. We have performed extensive experiments on side-scan sonar image with different targets and scales. The experimental results show that the detection accuracy of the proposed method reaches 95.73%, which is outperforms other state-of-the-art methods on the side-scan sonar image segmentation tasks. Moreover, the method has fewer computational parameters, facilitating future deployment it to underwater mobile detection devices.  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
25  
被引频次合计:  
25  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:000770054800116  
Indexed Date:  
2022-04-05  
Institution Address:  
Air Force Engn Univ, Sch Informat & Nav, Xian 710082, Shanxi, Peoples R China  
Institution Address:  
Xijing Univ, Sch Informat Engn, Xian 710123, Peoples R China  
Institution:  
Institution:

### 230. Multi-Scale Feature Fusion and Structure-Preserving Network for Face

摘要: Deep convolutional neural networks have demonstrated significant performance improvements in face super-resolution tasks. However, many deep learning-based approaches tend to overlook the inherent structural information and feature correlation across different scales in face images, making the accurate recovery of face structure in low-resolution cases challenging. To address this, this paper proposes a method that fuses multi-scale features while preserving the facial structure. It introduces a novel multi-scale residual block (MSRB) to reconstruct key facial parts and structures from spatial and channel dimensions, and utilizes pyramid attention (PA) to exploit non-local self-similarity, improving the details of the reconstructed face. Feature Enhancement Modules (FEM) are employed in the upscale stage to refine and enhance current features using multi-scale features from previous stages. The experimental results on CelebA, Helen and LFW datasets provide evidence that our method achieves superior quantitative metrics compared to the baseline, the Peak Signal-to-Noise Ratio (PSNR) outperforms the baseline by 0.282 dB, 0.343 dB, and 0.336 dB. Furthermore, our method demonstrates improved visual performance on two additional no-reference datasets, Widerface and Webface.  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
2076-3417  
入藏号:  
WOS:001046154800001  
Indexed Date:  
2023-08-25  
Institution Address:  
Hunan Normal Univ, Coll Informat Sci & Engn, Changsha 410081, Peoples R China  
Institution Address:  
Hunan Normal Univ, Sch Math & Stat, Changsha 410081, Peoples R China  
Institution Address:  
Changsha Univ Sci & Technol, Sch Comp & Commun Engn, Changsha 410114, Peoples R China  
Institution:  
Institution:  
Institution:

### 231. 被撤回的出版物: Evaluation of College Students' Ideological and Political

摘要: The construction of a correct worldview, outlook on life, and values for students is linked to the development and breakthrough in the management of ideological as well as political education of students. At the same time, college students must be encouraged to follow well-rounded education and struggle to be well-prepared for the challenges of the new era. In order to raise students' understanding of the critical role that political and ideological education plays in their academic success, it is authoritative that efforts to integrate these two spheres of learning be extended and new encounters made. That's what prompted this study, which is focused on assessing college students' level of ideological and political education administration, and it uses a mixture of big data technologies as well as artificial intelligence (AI) to do it. The accuracy of the traditional ideological as well as political education management quality assessment algorithm is not high, feature information extracted by the single-scale neural network (NN) is not rich enough, and the multiscale convolutional network (CN) fusion cannot consider the different values and importance for each scale. In this paper, the convolution kernel of the two-dimensional CN is changed to a one-dimensional convolution kernel, and the multiscale feature fusion CN model MCNN is first designed. The model is optimized and improved, the attention mechanism is integrated, and the MACNN model for the management evaluation of ideological as well as political education is proposed. Besides, this work organizes the network model in a wireless network environment that users can contact and operate at any time.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1939-0114  
eISSN:  
1939-0122  
入藏号:  
WOS:000805157000008  
Indexed Date:  
2022-06-13  
Institution Address:  
Tianjin Univ, Art Educ Ctr, Tianjin 300350, Peoples R China  
Institution Address:  
Tianjin Univ, Coll Management & Econ, Tianjin 300073, Peoples R China  
Institution:  
Institution:

### 232. Full Semantic Constructed Network for Urban Use Classification From Very

摘要: Recently, semantic segmentation technology has been a research hotspot in optical remote sensing urban use classification. However, because of coupled semantic relations in very high-resolution and complex urban scenes, a more effective semantic description for pixelwise urban use interpretation has become a challenge. Then, aiming to set up a more effective semantic description, the effective receptive field (ERF) is analyzed in general convolutional neural networks. The unreasonable ERF distribution in the stacked convolutional layers of the encoder would lead to a large amound of small ERFs and fewer not large enough ERFs that form a naive semantic description in decoder. Therefore, in this article, a novel full semantic constructed network (FSCNet) is proposed to improve the naive semantic description and set up an effective semantic description. First, to avoid noise from shallow feature layers, a residual refinement convolution is designed to optimize the full-scale skip connections based on the U-shaped encoder-decoder. Second, an interscale fusion module is newly designed for multiscale feature fusion, which can generate three initial semantic modalities that are prepared for redefining the full semantic description. Third, a multiscale local context spatial attention module and boundary supervision are designed for an initial shallow semantic modality to capture the pure boundary information, and then, pyramid spatial pooling is employed for an initial deep semantic modality to further enlarge the ERF and obtain more abstract global information. Next, a self-calibration convolution combined with the atrous spatial pyramid pooling is designed to rectify and enrich an initial middle semantic modality, which can improve the naive semantic description and bridge the semantic gap between the redefined shallow and deep semantic modalities to advance the full semantic feature fusion. Finally, extensive experiments are carried out on three benchmarks (e.g., ISPRS Vaihingen, Potsdam, and DLRSD), and comparative results show that the proposed FSCNet can get remarkable performance compared to state-of-the-art (SOTA) methods. Besides, the code is available at https://github.com/DorisCV/FSCNet.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000976155700004  
Indexed Date:  
2023-05-28  
Institution Address:  
Beijing Inst Technol, Beijing Key Lab Embedded Real Time Informat Proc T, Beijing 100081, Peoples R China  
Institution Address:  
Peking Univ, Sch Elect Engn & Comp Sci, Beijing 100871, Peoples R China  
Institution:  
Institution:

### 233. A Multiscale Instance Segmentation Method Based on Cleaning Rubber Ball

摘要: The identification of wear rubber balls in the rubber ball cleaning system in heat exchange equipment directly affects the descaling efficiency. For the problem that the rubber ball image contains impurities and bubbles and the segmentation is low in real time, a multi-scale feature fusion real-time instance segmentation model based on the attention mechanism is proposed for the object segmentation of the rubber ball images. First, we introduce the Pyramid Vision Transformer instead of the convolution module in the backbone network and use the spatial-reduction attention layer of the transformer to improve the feature extraction ability across scales and spatial reduction to reduce computational cost; Second, we improve the feature fusion module to fuse image features across scales, combined with an attention mechanism to enhance the output feature representation; Third, the prediction head separates the mask branches separately. Combined with dynamic convolution, it improves the accuracy of the mask coefficients and increases the number of upsampling layers. It also connects the penultimate layer with the second layer feature map to achieve detection of smaller images with larger feature maps to improve the accuracy. Through the validation of the produced rubber ball dataset, the Dice score, Jaccard coefficient, and mAP of the actual segmented region of this network with the rubber ball dataset are improved by 4.5%, 4.7%, and 7.73%, respectively, and our model achieves 33.6 fps segmentation speed and 79.3% segmentation accuracy. Meanwhile, the average precision of Box and Mask can also meet the requirements under different IOU thresholds. We compared the DeepMask, Mask R-CNN, BlendMask, SOLOv1 and SOLOv2 instance segmentation networks with this model in terms of training accuracy and segmentation speed and obtained good results. The proposed modules can work together to better handle object details and achieve better segmentation performance.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:000988067000001  
PubMed ID:  
37177464  
Indexed Date:  
2023-06-02  
Institution Address:  
Zhengzhou Univ, Sch Phys & Microelect, Zhengzhou 450001, Peoples R China  
Institution Address:  
Zhengzhou Univ, Inst Optoelect Informat Sci, Zhengzhou, Peoples R China  
Institution:  
Institution:

### 234. 3cDe-Net: a cervical cancer cell detection network based on an improved

摘要: Background Cervical cancer cell detection is an essential means of cervical cancer screening. However, for thin-prep cytology test (TCT)-based images, the detection accuracies of traditional computer-aided detection algorithms are typically low due to the overlapping of cells with blurred cytoplasmic boundaries. Some typical deep learning-based detection methods, e.g., ResNets and Inception-V3, are not always efficient for cervical images due to the differences between cervical cancer cell images and natural images. As a result, these traditional networks are difficult to directly apply to the clinical practice of cervical cancer screening. Method We propose a cervical cancer cell detection network (3cDe-Net) based on an improved backbone network and multiscale feature fusion; the proposed network consists of the backbone network and a detection head. In the backbone network, a dilated convolution and a group convolution are introduced to improve the resolution and expression ability of the model. In the detection head, multiscale features are obtained based on a feature pyramid fusion network to ensure the accurate capture of small cells; then, based on the Faster region-based convolutional neural network (R-CNN), adaptive cervical cancer cell anchors are generated via unsupervised clustering. Furthermore, a new balanced L1-based loss function is defined, which reduces the unbalanced sample contribution loss. Result Baselines including ResNet-50, ResNet-101, Inception-v3, ResNet-152 and the feature concatenation network are used on two different datasets (the Data-T and Herlev datasets), and the final quantitative results show the effectiveness of the proposed dilated convolution ResNet (DC-ResNet) backbone network. Furthermore, experiments conducted on both datasets show that the proposed 3cDe-Net, based on the optimal anchors, the defined new loss function, and DC-ResNet, outperforms existing methods and achieves a mean average precision (mAP) of 50.4%. By performing a horizontal comparison of the cells on an image, the category and location information of cancer cells can be obtained concurrently. Conclusion The proposed 3cDe-Net can detect cancer cells and their locations on multicell pictures. The model directly processes and analyses samples at the picture level rather than at the cellular level, which is more efficient. In clinical settings, the mechanical workloads of doctors can be reduced, and their focus can be placed on higher-level review work.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
1471-2342  
入藏号:  
WOS:000829164100001  
PubMed ID:  
35870877  
Indexed Date:  
2022-07-29  
Institution Address:  
Peking Univ Canc Hosp & Inst, Key Lab Carcinogenesis & Translat Res, Dept Gynecol Oncol, Minist Educ Beijing, Beijing 100142, Peoples R China  
Institution Address:  
Beijing Normal Univ, Sch Artificial Intelligence, Beijing 100875, Peoples R China  
Institution Address:  
Peking Univ Canc Hosp & Inst, Key Lab Carcinogenesis & Translat Res, Dept Pathol, Minist Educ Beijing, Beijing 100142, Peoples R China  
Institution:  
Peking Univ Canc Hosp & Inst  
Institution:  
Institution:  
Peking Univ Canc Hosp & Inst

### 235. Multiscale Feature Fusion for Hyperspectral Image Classification Using

摘要: Hyperspectral remote sensing images (HRSI) comprise three-dimensional image cubes, containing a single spectral dimension alongside two spatial dimensions. HRSI are presently among the foremost essential datasets for Earth observation. The task of HRSI classification is intricate due to the influence of spectral mixing, leading to notable variability within classes and resemblances across classes. Consequently, the field of HRSI classification has garnered significant research attention in recent times. Convolutional Neural Networks (CNNs) are harnessed to address these issues, enabling both feature extraction and classification. This study introduces a novel approach for HRSI classification called the hybrid 3D-2D depthwise separable convolution network (Hybrid DSCNet), which leverages multiscale feature integration. Within the Hybrid DSCNet, diverse kernel sizes contribute to an enriched feature extraction process from HRSI. The conventional 3D-2D CNN, while effective, comes with a computational load. Instead of using the standard 3D-2D CNN, this study adopts the 3D-2D DSC architecture. This approach partitions the conventional convolution into two components: pointwise and depthwise convolution, yielding a substantial reduction in trainable parameters and computational complexity. To evaluate the proposed method, the Indian Pines dataset along with WHU-Hi subdatasets (LongKou-LK, HanChuan-HC, and HongHu-HH) were employed. Employing a 5% training sample, impressive overall accuracy scores were achieved: 94.51%, 99.78%, 97.06%, and 97.27% for Indian Pines, WHU-LK, WHU-HC, and WHU-HH, respectively. Comparative analysis of the proposed approach with cutting-edge techniques within the literature reveals its superior performance across the four HRSI datasets. Notably, the Hybrid DSCNet attains enhanced classification accuracy while maintaining lower computational overhead.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0765-0019  
eISSN:  
1958-5608  
入藏号:  
WOS:001094288100012  
Indexed Date:  
2023-11-27  
Institution Address:  
Dicle Univ, Vocat Sch Tech Sci, TR-21200 Diyarbakir, Turkiye  
Institution Address:  
Harran Univ, Dept Comp Engn, TR-63050 Sanliurfa, Turkiye  
Institution Address:  
Harran Univ, Engn Fac, Dept Elect & Elect Engn, TR-63050 Sanliurfa, Turkiye  
Institution Address:  
Inonu Univ, Dept Comp Engn, TR-44280 Malatya, Turkiye  
Institution:  
Institution:  
Institution:  
Institution:

### 236. DBMF: Dual Branch Multiscale Feature Fusion Network for polyp

摘要: Accurate and reliable segmentation of colorectal polyps is important for the diagnosis and treatment of col-orectal cancer. Most of the existing polyp segmentation methods innovatively combine CNN with Transformer. Due to the single combination approach, there are limitations in establishing connections between local feature information and utilizing global contextual information captured by Transformer. Still not a better solution to the problems in polyp segmentation. In this paper, we propose a Dual Branch Multiscale Feature Fusion Network for Polyp Segmentation, abbreviated as DBMF, for polyp segmentation to achieve accurate segmentation of polyps. DBMF uses CNN and Transformer in parallel to extract multi-scale local information and global contextual information respectively, with different regions and levels of information to make the network more accurate in identifying polyps and their surrounding tissues. Feature Super Decoder (FSD) fuses multi-level local features and global contextual information in dual branches to fully exploit the potential of combining CNN and Transformer to improve the network's ability to parse complex scenes and the detection rate of tiny polyps. The FSD generates an initial segmentation map to guide the second parallel decoder (SPD) to refine the segmentation boundary layer by layer. SPD consists of a multi-scale feature aggregation module (MFA) and parallel polarized self-attention (PSA) and reverse attention fusion modules (RAF). MFA aggregates multi-level local feature information extracted by CNN Brach to find consensus regions between multiple scales and improve the network's ability to identify polyp regions. PSA uses dual attention to enhance the fine-grained nature of segmented regions and reduce the redundancy introduced by MFA and interference information. RAF mines boundary cues and establishes relationships between regions and boundary cues. The three RAFs guide the network to explore lost targets and boundaries in a bottom-up manner. We used the CVC-ClinicDB, Kvasir, CVC-300, CVC-ColonDB, and ETIS datasets to conduct comparison experiments and ablation experiments between DBMF and mainstream polyp segmentation networks. The results showed that DBMF outperformed the current mainstream networks on five benchmark datasets.  
Web of Science 核心合集中的 "被引频次":  
19  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
4  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
19  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:000900240200003  
PubMed ID:  
36401969  
Indexed Date:  
2023-01-06  
Institution Address:  
Shandong Technol & Business Univ, Coinnovat Ctr Shandong Coll & Univ Future Intellig, Sch Informat & Elect Engn, Yantai 264005, Peoples R China  
Institution Address:  
Shandong Univ Finance & Econ, Sch Comp Sci & Technol, Jinan 250014, Shandong, Peoples R China  
Institution:  
Institution:

### 237. MBUTransNet: multi-branch U-shaped network fusion transformer

摘要: PurposeRecently, transformers have been adopted to computer vision applications and achieve great success in image segmentation. However by simply applying transformers to medical segmentation task it is hard to achieve much higher accuracy than by traditional U-shaped network structures, which are based on CNNs and has been extensively researched. On the other hand, CNN structure pays more attention to local information and ignores global information, which is very important for the medical image segmentation dataset with cell scattered background. This motivates us to explore the feasibility of using U-shape effective fusion transformer network architectures for medical image segmentation tasks.MethodsIn this paper, we propose a multibranch U-shaped structure fusion transformer network (MBUTransNet), which consists of two distinct branches. In branch 1, Coordinate attention transformer is designed to extract long-term dependency information through weight coordinates. In branch 2, small U-net blocks and multiscale feature fusion block are proposed to replace convolution blocks of each layer and fuse the feature maps from different layers, respectively.ResultsOur experiments demonstrate that the proposed MBUTransNet has achieved a 0.076 and 0.1269 improvement in DICE compared to the previous best method on MoNuSeg and Synapse multiorgan segmentation dataset, respectively, while the model parameters will be no significant increase.ConclusionWithout bells and whistles, MBUTransNet achieves better performance on medical image datasets, including medical cell segmentation and abdominal organs segmentation. Compared with transformer-based methods, our proposed model also obtains quite competitive parameters.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
4  
ISSN:  
1861-6410  
eISSN:  
1861-6429  
入藏号:  
WOS:000963958900001  
PubMed ID:  
37024630  
Indexed Date:  
2023-04-24  
Institution Address:  
Liaoning Tech Univ, Sch Elect & Informat Engn, 188 Longwan South St, Huludao 125105, Liaoning, Peoples R China  
Institution Address:  
Linyi Univ, Sch Informat Sci & Engn, Shuangling Rd, Linyi 276000, Shandong, Peoples R China  
Institution:  
Institution:

### 238. Attention-enhanced multi-scale residual network for single image

摘要: Single image super-resolution (SISR) has important applications in many fields. With the help of this technology, the broadband requirement of image transmission can be reduced, the effect of remote sensing observation can be improved, and the location of lesion cells can be accurately located. Convolutional neural networks (CNNs) using multi-scale feature extraction structure can gain a large amount of information from a low-resolution input, which is helpful to improve the performance of SISR. However, these CNNs usually treat different types of information equally. There is a lot of redundancy in the information obtained, which limits the representation ability of the networks. We proposed an attention-enhanced multi-scale residual block (AMRB), which increases the proportion of useful information by embedding convolutional block attention module. Furthermore, we construct an attention-enhanced multi-scale residual network based on one time feature fusion (OAMRN). Extensive experiments illustrate the necessity of the AMRB and the superiority of proposed OAMRN over the state-of-the-art methods in terms of both quantitative metrics and visual quality.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
8  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:000743447200002  
Indexed Date:  
2022-03-03  
Institution Address:  
Southwest Univ, Coll Elect & Informat Engn, Chongqing Key Lab Nonlinear Circuit & Intelligent, Chongqing 400715, Peoples R China  
Institution Address:  
Chongqing Univ Sci & Technol, Coll Chem & Chem Engn, 20 East Daxuecheng Rd, Chongqing 401331, Peoples R China  
Institution:  
Institution:

### 239. DB-Net: dual-branch deep learning network for cloud detection utilizing

摘要: Remote sensing images often contain a significant amount of clouds, which can result in substantial resource costs during transmission and storage. Cloud detection can reduce these costs. Although current cloud detection methods perform well in extracting large and thick clouds, there are still some issues, such as missed detection of small and thin clouds and false detection in non-cloud areas. Therefore, we propose a deep learning framework called DB-Net. It consists of three main modules: feature extraction module (FEM), cascaded feature enhancement module (CFEM), and feature fusion module (FFM). In the FEM, we leverage the advantages of both convolutional neural network and Transformer by utilizing two branches to reduce the loss of semantic information. To enhance the acquisition capability of multi-scale semantic information, in the CFEM, regular convolutions are replaced with deformable convolutions to adaptively capture cloud features of various sizes, and a cascaded structure is designed to enhance the interaction of information among different scales. Furthermore, to focus on small and thin cloud information and suppress non-cloud background information, we designed the FFM using attention mechanisms to enhance the target information in the features extracted by FEM and CFEM. Extensive experiments were conducted on the GF1-WHU dataset, and comparisons were made with mainstream cloud detection networks. The experimental results indicate that the proposed DB-Net method reduces cloud information omission, effectively focuses on thin clouds and small clouds, and improves overall cloud detection performance.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1931-3195  
入藏号:  
WOS:001330379800017  
Indexed Date:  
2024-10-18  
Institution Address:  
Shanghai Ocean Univ, Sch Informat, Remote Sensing Nav Lab, Shanghai, Peoples R China  
Institution Address:  
Sense Time Intelligent Technol Co Ltd, Shanghai, Peoples R China  
Institution:  
Institution:  
Sense Time Intelligent Technol Co Ltd

### 240. MFRNet: A Multipath Feature Refinement Network for Semantic Segmentation

摘要: Deep convolutional neural networks have made significant progress in the field of intelligent analysis of remote-sensing images. However, the semantic segmentation task in high-resolution remote-sensing (HRRS) images always faces the problem of large-scale variation and complex background samples, which causes difficulties in distinguishing confusable ground objects. In this letter, we propose a novel multipath feature refinement network (MFRNet) to alleviate the above problems. We design the feature refinement module (FRM) to fuse features at various scales, which helps to capture different levels of spatial information. It also alleviates the boundary ambiguity problem by enhancing the learning of features with boundary information. The multiscale feature attention module (MFAM) combines atrous convolution and non-local block to obtain larger receptive fields and long-range contextual information, while the feature fusion module (FFM) balances semantic and spatial information, further improving the embedding of locally discriminative features. Experimental results on ISPRS Potsdam and LoveDA datasets indicate that the proposed MFRNet outperforms other semantic segmentation methods and excels in the accuracy and consistency of object boundary segmentation.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
2150-704X  
eISSN:  
2150-7058  
入藏号:  
WOS:000889106700001  
Indexed Date:  
2022-12-04  
Institution Address:  
Shandong Univ, Sch Software, Jinan, Shandong, Peoples R China  
Institution Address:  
Heze Univ, Sch Comp, Heze, Peoples R China  
Institution:  
Institution:

### 241. SAR Target Classification Based on Multiscale Attention Super-Class

摘要: The convolutional neural network (CNN) is widely used in synthetic aperture radar (SAR) target recognition, but conventional CNN mainly adopts a single-scale convolutional kernel, resulting in losing part of the feature information of targets and does not pay enough attention to significant features. On the other hand, conventional CNN approaches only assign fine-class labels to SAR targets, ignoring the high-level semantics information of similar categories, which reduces the feature differences between categories and the generalization ability of the model. Therefore, this article proposes a multiscale attention super-class CNN (MSA-SCNN) for SAR target classification. First, MSA-SCNN combines multiscale feature fusion with the attention module to improve the integrity of SAR target feature representation. The attention module includes channel and spatial attention modules, which realize the weighted enhancement of different scale features. Additionally, MSA-SCNN introduces super-class labels to increase the feature difference between categories. The classification stage consists of a fine-class branch and a super-class branch, and the features trained on the super-class branch are fused to the fine-class branch to improve the network's fine classification ability. Experiments on the moving and stationary target acquisition and recognition dataset and the FUSAR-Ship dataset show that the proposed MSA-SCNN outperforms many current existing state-of-the-art methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
16  
被引频次合计:  
16  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000873830800008  
Indexed Date:  
2022-11-11  
Institution Address:  
Natl Univ Def Technol, Coll Elect Sci, Changsha 410073, Peoples R China  
Institution:

### 242. MHST: Multiscale Head Selection Transformer for Hyperspectral and LiDAR

摘要: The joint use of hyperspectral image (HSI) and light detection and ranging (LiDAR) data has gained significant performance on land-cover classification. Although spatial-spectral feature learning methods based on convolutional neural networks and transformer networks have achieved prominent advances, contextual information described by fixed convolutional kernels and all self-attention heads selected have limited ability to characterize the detailed information and nonredundant features of land-covers on multimodal data. In this article, a multiscale head selection transformer (MHST) network, is proposed to fully explore detailed and nonredundant features in spatial and spectral dimensions of HSI and LiDAR data. To better acquire detailed information of spatial and spectral features at different scales, a multiscale spectral-spatial feature extraction module, including cascaded multiscale 3-D and 2-D convolutional layers, is inserted into MHST. Simultaneously, an adaptive global feature extraction module based on head selection pooling transformer is given after transformer encoder module for alleviating token redundancy in an adaptive computation style. Finally, we develop a multimodal-multiscale feature fusion classification module with local features and global class token, to exploit a powerful global-local fuse style. The extensive experiments on three popular datasets demonstrate that MHST significantly outperforms other related networks.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
被引频次合计:  
8  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001180713400013  
Indexed Date:  
2024-04-02  
Institution Address:  
Nanjing Univ Posts & Telecommun, Sch Comp Sci, Nanjing 210023, Peoples R China  
Institution Address:  
Jiangsu Prov Engn Res Ctr Airborne Detecting & In, Coll Automation, Coll Artificial Intelligence, Nanjing 210049, Peoples R China  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Minist Educ, Key Lab Radar Imaging & Microwave Photon, Nanjing 211106, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Coll Automat, Nanjing 210023, Peoples R China  
Institution Address:  
Nanjing Univ Posts & Telecommun, Coll Artificial Intelligence, Nanjing 210023, Peoples R China  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Key Lab Radar Imaging & Microwave Photon, Minist Educ, Nanjing, Peoples R China  
Institution:  
Institution:  
Jiangsu Prov Engn Res Ctr Airborne Detecting & In  
Institution:  
Institution:  
Institution:  
Institution:

### 243. A Hybrid Algorithm with Swin Transformer and Convolution for Cloud

摘要: Cloud detection is critical in remote sensing image processing, and convolutional neural networks (CNNs) have significantly advanced this field. However, traditional CNNs primarily focus on extracting local features, which can be challenging for cloud detection due to the variability in the size, shape, and boundaries of clouds. To address this limitation, we propose a hybrid Swin transformer-CNN cloud detection (STCCD) network that combines the strengths of both architectures. The STCCD network employs a novel dual-stream encoder that integrates Swin transformer and CNN blocks. Swin transformers can capture global context features more effectively than traditional CNNs, while CNNs excel at extracting local features. The two streams are fused via a fusion coupling module (FCM) to produce a richer representation of the input image. To further enhance the network's ability in extracting cloud features, we incorporate a feature fusion module based on the attention mechanism (FFMAM) and an aggregation multiscale feature module (AMSFM). The FFMAM selectively merges global and local features based on their importance, while the AMSFM aggregates feature maps from different spatial scales to obtain a more comprehensive representation of the cloud mask. We evaluated the STCCD network on three challenging cloud detection datasets (GF1-WHU, SPARCS, and AIR-CD), as well as the L8-Biome dataset to assess its generalization capability. The results show that the STCCD network outperformed other state-of-the-art methods on all datasets. Notably, the STCCD model, trained on only four bands (visible and near-infrared) of the GF1-WHU dataset, outperformed the official Landsat-8 Fmask algorithm in the L8-Biome dataset, which uses additional bands (shortwave infrared, cirrus, and thermal).  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
14  
eISSN:  
2072-4292  
入藏号:  
WOS:001100318600001  
Indexed Date:  
2023-12-05  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst AIR, Beijing 100094, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 244. Three-dimensional hybrid fusion networks for current-based bearing fault

摘要: Intelligent fault diagnosis (IFD) techniques commonly use vibration-based measurements to perform health monitoring of critical rotating components in industrial systems. However, these vibration-based approaches may be limited in cost-sensitive applications, because the installation of vibration sensors is inconvenient and vibration sensors are expensive. Considering the difficulties of IFD using only current-related information from the motor current signal (MCS), this paper proposes a three-dimensional hybrid-fusion neural network (3D-HFN) that can automatically perform both data- and feature-level fusion of multi-phase current signals for MCS-based IFD of the rolling bearing. The 3D-HFN consists of the multivariate variational mode decomposition (MVMD) and an improved three-dimensional convolution neural network (3D-CNN). Firstly, MVMD is proposed to process multi-phase current signals, which adaptively acquire several intrinsic mode functions with mode-alignment properties. Subsequently, signal-to-image conversion and 3D stacking methods are used to construct 3D-like data in the current-phase dimension, which can fully preserve the interaction relationship between different phases using data-level fusion. Finally, an improved 3D-CNN with multiscale feature fusion and the smooth maximum unit is proposed to learn the 3D-like data and identify different health conditions for the rolling bearing. An open-source dataset with composite bearing faults is used to validate the merits of the proposed method. Experimental results show that the proposed approach has achieved more reliable diagnosis performance than other hand-crafted or 2D/3D-CNN-based algorithms in MCS-based IFD of the rolling bearing.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001104165100001  
Indexed Date:  
2023-12-17  
Institution Address:  
Huazhong Univ Sci & Technol, Sch Aerosp Engn, Wuhan 430074, Peoples R China  
Institution Address:  
Georgia Inst Technol, George W Woodruff Sch Mech Engn, Atlanta, GA 30332 USA  
Institution:  
Institution:

### 245. FSNB-YOLOV8: Improvement of Object Detection Model for Surface Defects

摘要: The current object detection algorithm based on CNN makes it difficult to effectively capture the characteristics of subtle defects in online industrial product packaging bags. These defects are often visually similar to the texture or background of normal product packaging bags, and the model cannot effectively distinguish them. In order to deal with these challenges, this paper optimizes and improves the network structure based on YOLOv8 to achieve accurate identification of defects. First, in order to solve the long-tail distribution problem of data, a fuzzy search data enhancement algorithm is introduced to effectively increase the number of samples. Secondly, a joint network of FasterNet and SPD-Conv is proposed to replace the original backbone network of YOLOv8, which effectively reduces the computing load and improves the accuracy of defect identification. In addition, in order to further improve the performance of multiscale feature fusion, a weighted bidirectional feature pyramid network (BiFPN) is introduced, which effectively enhances the model's ability to detect defects at different scales through the fusion of deep information and shallow information. Finally, in order to reduce the sensitivity of the defect position deviation, the NWD loss function is used to optimize the positioning performance of the model better and reduce detection errors caused by position errors. Experimental results show that the FSNB\_YOLOv8 model proposed in this paper can reach 98.8% mAP50 accuracy. This success not only verifies the effectiveness of the optimization and improvement of this article's model but also provides an efficient and accurate solution for surface defect detection of industrial product packaging bags on artificial assembly systems.  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001311008600001  
Indexed Date:  
2024-09-16  
Institution Address:  
Beijing Informat Sci & Technol Univ, Sch Comp Sci, Beijing 100192, Peoples R China  
Institution Address:  
Natl Inst Biol Sci, Beijing 102206, Peoples R China  
Institution Address:  
Natl Comp Syst Engn Res Inst China, Beijing 100083, Peoples R China  
Institution:  
Institution:  
Institution:  
Natl Comp Syst Engn Res Inst China

### 246. CRGF-YOLO: An Optimized Multi-Scale Feature Fusion Model Based on YOLOv5

摘要: The identification of imperfections on steel surfaces is vital for ensuring the quality of industrial products. It requires the capability of real-time detection with high accuracy. This paper proposes the CRGF-YOLO (Contextual Reparameterized Generalized Feature) model based on YOLOv5. In the network, BottleneckCSP structures and depthwise separable convolutions utilizing the structural reparameterization are introduced to reduce the model size and improve performance. In addition, contextual transformer modules are employed as self-attention mechanisms to improve feature representations by capturing long-range dependencies, outperforming conventional convolutional networks. Furthermore, the simplified generalized feature pyramid network is embedded to aggregate multi-scale feature maps and enhance the network's robustness. Finally, four prediction heads with different sizes are employed to predict defects, which are supported by prior bounding boxes generated using k-means clustering algorithm. The Focal-EIOU (Exponential Intersection over Union) loss function is introduced to improve detection accuracy and expedite model convergence. The improved model achieves a mean average precision (mAP) of 82.2% on the NEU-DET dataset, outperforming the baseline YOLOv5s by 7.7% mAP while maintaining real-time speeds. Comparative evaluations demonstrate CRGF-YOLO's superior performance over previous state-of-the-art methods like Faster R-CNN (77.4% mAP), YOLOv3 (77.4% mAP), YOLOv7s (72.1% mAP), and YOLOv8s (78.7% mAP) for steel surface defect detection. Overall, this study provides valuable insights and practical guidance for the advancement of defect detection technology.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1875-6891  
eISSN:  
1875-6883  
入藏号:  
WOS:001250995000001  
Indexed Date:  
2024-06-25  
Institution Address:  
Natl Univ Def Technol, Coll Intelligence Sci & Technol, Changsha 410073, Peoples R China  
Institution Address:  
Natl Univ Def Technol, Lab Sci & Technol Integrated Logist Support, Changsha 410073, Peoples R China  
Institution:  
Institution:

### 247. Multi-scale feature extraction and TrasMLP encoder module for ocean HABs

摘要: Due to tiny edge and texture details of harmful algae blooms(HABs), existing segmentation networks are not effective for HABs segmentation. In order to solve the above problems, this paper proposes a Multi-scale Feature extraction and TrasMLP Encoder Fusion-based network (MFTS). To tackle the complex morphological characteristics and the complex backgrounds of HABs, a TrasMlp module which can effectively identify long-range patterns and adapt network parameters is introduced, enabling accurately parsing of complex algae images. Secondly, the deep convolution module is constructed by combining deep separable convolution with a twochannel attention mechanism to separate the target region from the background. In addition, this paper proposes a Weighted Feature Fusion of Deep Convolution and INRS Encoder Module classification network(FDIR) is proposed to quantify the performance of the image segmentation network. The segmentation results on HABs dataset from AICO Lab show that our proposed MFTS model achieves a miou of 90.02%, outperforming the performance of classical segmentation networks such as U-Net and Mask R-CNN. Compared to original HABs dataset, the segmented result shows a 5.1% improvement in the classification accuracy of the FDIR model.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0029-8018  
eISSN:  
1873-5258  
入藏号:  
WOS:001293576100001  
Indexed Date:  
2024-08-23  
Institution Address:  
Shandong Univ Sci & Technol, Coll Comp Sci & Engn, Qingdao 266590, Peoples R China  
Institution:

### 248. A MFFBSNet crowd counting algorithm based on multi-scale feature fusion

摘要: Aiming at the problems of scale variation, uneven distribution, and background occlusion of dense crowds in complex scenes, a crowd counting algorithm MFFBSNet based on multi-scale feature fusion and background suppression is proposed.The first 13 layers of the visual geometry group network VGG-16 are utilized as the front-end of the network. An atrous spatial pyramid pooling (ASPP) and a pyramid split attention (PSA) mechanism based on a lightweight design are introduced to construct a multi-scale feature fusion module, which addresses the problem of scale variation in dense crowds; In the middle of this network, spatial and channel attention mechanisms are incorporated to refine the feature maps, highlighting the head regions in the image; The backend of this network employs atrous convolution, which enlarges the receptive field without losing image resolution, to generate a background segmentation attention map. This suppresses background noise in the image and enhances the quality of the crowd density map. Experimental results on three public datasets, namely ShanghaiTech, UCF\_CC\_50, and NWPU-Crowd,demonstrate that the proposed crowd counting algorithm based on the MFFBSNet achieves higher counting accuracy compared to methods such as MCNN,SwitchCNN,and CSRNet.  
摘要:  
针对复杂场景中的密集人群尺度变化、分布不均匀、背景遮挡等问题,提出一种基于多尺度特征融合与背景抑制的MFFBSNet人群计数算法。以视觉几何组网络VGG-16的前13层作为网络前端部分,引入空洞空间卷积池化金字塔(ASPP)和基于轻量级金字塔切分注意力机制(PSA)构建多尺度特征融合模块,以解决密集人群尺度变化问题;在网络的中间部分加入空间注意力机制以及通道注意力机制对特征图进行校准,突出图像人头区域;网络后端部分使用可加大感受野且不丢失图像分辨率的空洞卷积生成背景分割注意力图,抑制图像中背景噪声,提升人群分布密度图的质量。在ShanghaiTech、UCF\_CC\_50及NWPU-Crowd 3个公开数据集上的实验结果表明,相较于MCNN、SwitchCNN、CSRNet等算法,提出的基于MFFBSNet的人群计数算法的计数准确度较高。  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1007-130X  
入藏号:  
CSCD:7868162  
Indexed Date:  
2025-03-01  
Institution Address:  
浙江师范大学计算机科学与技术学院(人工智能学院);;嘉兴市智慧交通重点实验室, ;;嘉兴市智慧交通重点实验室, 金华;;嘉兴, 浙江;;浙江 321004;;314001, 中国  
Institution Address:  
浙江师范大学计算机科学与技术学院(人工智能学院), 金华, 浙江 321004, 中国  
Institution Address:  
嘉兴市智慧交通重点实验室;;嘉兴南湖学院信息工程学院;;嘉兴市智能计算与数据科学重点实验室, 嘉兴市智慧交通重点实验室;;;;嘉兴市智能计算与数据科学重点实验室, 嘉兴;;嘉兴;;嘉兴, 浙江;;浙江;;浙江 314001;;314001;;314001, 中国  
Institution Address:  
嘉兴市智慧交通重点实验室;;嘉兴大学信息科学与工程学院, 嘉兴市智慧交通重点实验室;;, 嘉兴;;嘉兴, 浙江;;浙江 314001;;314001, 中国  
Institution Address:  
School of Computer Science and Technology(School of Artificial Intelligence), Zhejiang Normal University;;Jiaxing Key Laboratory of Smart Transportations, ;;Jiaxing Key Laboratory of Smart Transportations, Jinhua;;Jiaxing, ;; 321004;;314001  
Institution Address:  
School of Computer Science and Technology(School of Artificial Intelligence), Zhejiang Normal University, Jinhua, Zhejiang 321004, China  
Institution Address:  
Jiaxing Key Laboratory of Smart Transportations;;College of Information Engineering, Jiaxing Nanhu University;;Jiaxing Key Laboratory of Intelligent Computation and Data Science, Jiaxing Key Laboratory of Smart Transportations;;;;Jiaxing Key Laboratory of Intelligent Computation and Data Science, Jiaxing;;Jiaxing;;Jiaxing, ;;;; 314001;;314001;;314001  
Institution Address:  
Jiaxing Key Laboratory of Smart Transportations;;College of Information Science and Engineering, Jiaxing University, Jiaxing Key Laboratory of Smart Transportations;;, Jiaxing;;Jiaxing, ;; 314001;;314001  
Institution:  
浙江师范大学计算机科学与技术学院(人工智能学院);;嘉兴市智慧交通重点实验室  
Institution:  
浙江师范大学计算机科学与技术学院(人工智能学院)  
Institution:  
嘉兴市智慧交通重点实验室;;嘉兴南湖学院信息工程学院;;嘉兴市智能计算与数据科学重点实验室  
Institution:  
嘉兴市智慧交通重点实验室;;嘉兴大学信息科学与工程学院  
Institution:  
School of Computer Science and Technology(School of Artificial Intelligence), Zhejiang Normal University;;Jiaxing Key Laboratory of Smart Transportations  
Institution:  
School of Computer Science and Technology(School of Artificial Intelligence), Zhejiang Normal University  
Institution:  
Jiaxing Key Laboratory of Smart Transportations;;College of Information Engineering, Jiaxing Nanhu University;;Jiaxing Key Laboratory of Intelligent Computation and Data Science  
Institution:  
Jiaxing Key Laboratory of Smart Transportations;;College of Information Science and Engineering, Jiaxing University

### 249. Towards Accurate Oriented Object Detection in Aerial Images with

摘要: Detecting objects in aerial images is a long-standing and challenging problem since the objects in aerial images vary dramatically in size and orientation. Most existing neural network based methods are not robust enough to provide accurate oriented object detection results in aerial images since they do not consider the correlations between different levels and scales of features. In this paper, we propose a novel two-stage network-based detector with adaptive feature fusion towards highly accurate oriented object detection in aerial images, named AFF-Det. First, a multi-scale feature fusion module (MSFF) is built on the top layer of the extracted feature pyramids to mitigate the semantic information loss in the small-scale features. We also propose a cascaded oriented bounding box regression method to transform the horizontal proposals into oriented ones. Then the transformed proposals are assigned to all feature pyramid network (FPN) levels and aggregated by the weighted RoI feature aggregation (WRFA) module. The above modules can adaptively enhance the feature representations in different stages of the network based on the attention mechanism. Finally, a rotated decoupled-RCNN head is introduced to obtain the classification and localization results. Extensive experiments are conducted on the DOTA and HRSC2016 datasets to demonstrate the advantages of our proposed AFF-Det. The best detection results can achieve 80.73% mAP and 90.48% mAP, respectively, on these two datasets, outperforming recent state-of-the-art methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
16  
被引频次合计:  
16  
ISSN:  
1551-6857  
eISSN:  
1551-6865  
入藏号:  
WOS:000927167400006  
Indexed Date:  
2023-03-08  
Institution Address:  
Shanghai Jiao Tong Univ, 800 Dongchuan Rd, Shanghai 200240, Peoples R China  
Institution Address:  
Beijing Inst Astronaut Syst Engn, 1 Donggaodi South St, Beijing 100076, Peoples R China  
Institution:  
Institution:  
Beijing Inst Astronaut Syst Engn

### 250. A Lightweight Object Detection Framework for Remote Sensing Images

摘要: Onboard real-time object detection in remote sensing images is a crucial but challenging task in this computation-constrained scenario. This task not only requires the algorithm to yield excellent performance but also requests limited time and space complexity of the algorithm. However, previous convolutional neural networks (CNN) based object detectors for remote sensing images suffer from heavy computational cost, which hinders them from being deployed on satellites. Moreover, an onboard detector is desired to detect objects at vastly different scales. To address these issues, we proposed a lightweight one-stage multi-scale feature fusion detector called MSF-SNET for onboard real-time object detection of remote sensing images. Using lightweight SNET as the backbone network reduces the number of parameters and computational complexity. To strengthen the detection performance of small objects, three low-level features are extracted from the three stages of SNET respectively. In the detection part, another three convolutional layers are designed to further extract deep features with rich semantic information for large-scale object detection. To improve detection accuracy, the deep features and low-level features are fused to enhance the feature representation. Extensive experiments and comprehensive evaluations on the openly available NWPU VHR-10 dataset and DIOR dataset are conducted to evaluate the proposed method. Compared with other state-of-art detectors, the proposed detection framework has fewer parameters and calculations, while maintaining consistent accuracy.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
24  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
25  
eISSN:  
2072-4292  
入藏号:  
WOS:000624517000001  
Indexed Date:  
2021-03-31  
Institution Address:  
Northwestern Polytech Univ, Shaanxi Prov Key Lab Speech & Image Informat Proc, Natl Engn Lab Integrated Aerospace Ground Ocean B, Sch Comp Sci, Xian 710129, Peoples R China  
Institution Address:  
CAST Xian, Key Lab Sci & Technol Space Microwave, Xian 710100, Peoples R China  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Commun & Informat Engn, Xian 710121, Shaanxi, Peoples R China  
Institution:  
Institution:  
CAST Xian  
Institution:

### 251. Multi-attention semantic segmentation method for forest information

摘要: The southern hilly region of China boasts abundant forest resources, which are crucial for maintaining ecological stability. However, the complex vegetation structure and fragmented terrain in this area lead to intricate and disorderly forest types, resulting in semantic confusion among vegetation in remote sensing images. Consequently, accurately classifying forest types poses significant challenges. We propose a semantic segmentation model with multiple attention mechanisms using convolutional neural networks. We enhance the U-Net model's encoder with a deeper convolutional network to expand the receptive field without significant computation increase. Furthermore, we integrate spatial attention within the U-Net's skip connections and multiscale feature fusion. Experimentally, the multiple attention mechanism U-Net model outperforms the original, averaging 90.67% intersection over union, 94.33% pixel accuracy, and 96.00% classification accuracy for 0.5 m resolution forest type classification. These improvements are 8.00%, 4.33%, and 5.00%, respectively. The model accurately distinguishes forest types in the southern hilly region, enabling precise information-based forest supervision.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1931-3195  
入藏号:  
WOS:001270024600016  
Indexed Date:  
2024-07-24  
Institution Address:  
Jiangxi Univ Sci & Technol, Jiangxi Prov Key Lab Water Ecol Conservat Headwate, Ganzhou, Peoples R China  
Institution Address:  
Jiangxi Prov Geol Bur Geog Informat Engn Brigade, Geog Informat Engn Brigade, Nanchang, Peoples R China  
Institution Address:  
East China Univ Technol, Sch Surveying & Mapping Engn, Nanchang, Peoples R China  
Institution Address:  
Jiangxi Univ Sci & Technol, Sch Sci, Ganzhou, Peoples R China  
Institution:  
Institution:  
Jiangxi Prov Geol Bur Geog Informat Engn Brigade  
Institution:  
Institution:

### 252. False Data Injection Attack Detection Method Based on Deep Learning With

摘要: Cyber-attacks, especially the false data injection attack (FDIA), are gradually becoming a common way to threaten the regular operation of power grid. However, the FDIA is challenging to detect because it prevents the bad data detection mechanism in the energy management system from destroying the integrity of measurement information. Aiming at the problem of the FDIA detection in smart grids, this paper presents a FDIA detection method based on deep learning with multi-scale feature fusion. First, the improved convolution neural network (ICNN) is used to predict measurement data by combining convolution neural network with the Inception v1 module. Then, the attention mechanism is introduced into the ICNN to extract and fuse full and partial features of measurement data. By fitting the function between measurement and state vectors, the state data are generated with predicted measurement data. Eventually, the threshold of divergence is obtained to determine whether the FDIA occurs or not by the difference in probability distribution between predicted and actual state vectors. The performance of the proposed method is evaluated in the IEEE 14-node and 39-node test systems. The results show that the proposed method can accurately detect the existence of FDIA in time. This method has definite robustness to noise and distributed generation switching.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001262675200001  
Indexed Date:  
2024-07-14  
Institution Address:  
Shandong Univ Technol, Sch Elect & Elect Engn, Zibo 255000, Peoples R China  
Institution Address:  
Zibo Metrol Technol Res Inst, Zibo 255025, Peoples R China  
Institution:  
Institution:  
Zibo Metrol Technol Res Inst

### 253. TW-YOLO: An Innovative Blood Cell Detection Model Based on Multi-Scale

摘要: As deep learning technology has progressed, automated medical image analysis is becoming ever more crucial in clinical diagnosis. However, due to the diversity and complexity of blood cell images, traditional models still exhibit deficiencies in blood cell detection. To address blood cell detection, we developed the TW-YOLO approach, leveraging multi-scale feature fusion techniques. Firstly, traditional CNN (Convolutional Neural Network) convolution has poor recognition capabilities for certain blood cell features, so the RFAConv (Receptive Field Attention Convolution) module was incorporated into the backbone of the model to enhance its capacity to extract geometric characteristics from blood cells. At the same time, utilizing the feature pyramid architecture of YOLO (You Only Look Once), we enhanced the fusion of features at different scales by incorporating the CBAM (Convolutional Block Attention Module) in the detection head and the EMA (Efficient Multi-Scale Attention) module in the neck, thereby improving the recognition ability of blood cells. Additionally, to meet the specific needs of blood cell detection, we designed the PGI-Ghost (Programmable Gradient Information-Ghost) strategy to finely describe the gradient flow throughout the process of extracting features, further improving the model's effectiveness. Experiments on blood cell detection datasets such as BloodCell-Detection-Dataset (BCD) reveal that TW-YOLO outperforms other models by 2%, demonstrating excellent performance in the task of blood cell detection. In addition to advancing blood cell image analysis research, this work offers strong technical support for future automated medical diagnostics.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:001332827200001  
PubMed ID:  
39409208  
Indexed Date:  
2024-10-21  
Institution Address:  
Zhejiang Sci Tech Univ, Sch Comp Sci & Technol, Hangzhou 310018, Peoples R China  
Institution:

### 254. Three-Dimensional Reconstruction Methods for Obstacles in Complex

摘要: Detecting irregular obstacles under complex scenarios of intelligent parking is a difficult task. Therefore, a method that employs a gridded structured light projection for the detection area is proposed in this study. Specifically, this method captures the deformation of structured light grids on obstacle surfaces, thereby enhancing the precision of obstacle feature collection. In addition, a method for generating depth maps via the training of an end- to- end network is introduced. Subsequently, the fusion of external contour features from red green blue (RGB) images with three-dimensional (3D) depth features from depth images is achieved, culminating in the proposition of a dual- feature parallel processing algorithm for RGB and depth imagery. A multi- scale feature fusion extraction model is designed, facilitating multifaceted feature extraction and in-depth fusion without escalating model complexity, which enables the transition of mesh models towards accurate 3D representations. Consequently, a multi- scale feature- informed, graph convolutional neural network- based end- to- end 3D reconstruction model is established. Experimental results in intelligent parking scenarios indicate that compared to foundational 3D reconstruction models, the model proposed herein achieves a mean reduction of 2 degrees o and 9 degrees o in chamfer distance and earth mover's distance, respectively. Furthermore, relative to three mainstream 3D reconstruction models, the mean reduction in chamfer distance is 60 degrees o, 2 degrees o, and 78 degrees o, respectively, while the reduction in earth mover's distance is 16 degrees o, 23 degrees o, and 91 degrees o, respectively.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-4125  
入藏号:  
WOS:001397832800001  
Indexed Date:  
2025-01-22  
Institution Address:  
Jiangsu Univ, Automot Engn Res Inst, Zhenjiang 212013, Jiangsu, Peoples R China  
Institution Address:  
Jiangsu Univ, Sch Automot & Traff Engn, Zhenjiang 212013, Jiangsu, Peoples R China  
Institution Address:  
Jiangsu Univ, Sch Comp Sci & Commun Engn, Zhenjiang 212013, Jiangsu, Peoples R China  
Institution:  
Institution:  
Institution:

### 255. RockSeg: A Novel Semantic Segmentation Network Based on a Hybrid

摘要: Rock detection on the surface of celestial bodies is critical in the deep space environment for obstacle avoidance and path planning of space probes. However, in the remote and complex deep environment, rocks have the characteristics of irregular shape, being similar to the background, sparse pixel characteristics, and being easy for light and dust to affect. Most existing methods face significant challenges to attain high accuracy and low computational complexity in rock detection. In this paper, we propose a novel semantic segmentation network based on a hybrid framework combining CNN and transformer for deep space rock images, namely RockSeg. The network includes a multiscale low-level feature fusion (MSF) module and an efficient backbone network for feature extraction to achieve the effective segmentation of the rocks. Firstly, in the network encoder, we propose a new backbone network (Resnet-T) that combines the part of the Resnet backbone and the transformer block with a multi-headed attention mechanism to capture the global context information. Additionally, a simple and efficient multiscale feature fusion module is designed to fuse low-level features at different scales to generate richer and more detailed feature maps. In the network decoder, these feature maps are integrated with the output feature maps to obtain more precise semantic segmentation results. Finally, we conduct experiments on two deep space rock datasets: the MoonData and MarsData datasets. The experimental results demonstrate that the proposed model outperforms state-of-the-art rock detection algorithms under the conditions of low computational complexity and fast inference speed.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
5  
eISSN:  
2072-4292  
入藏号:  
WOS:001061369800001  
Indexed Date:  
2023-09-17  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Comp Sci & Technol, Nanjing 211106, Peoples R China  
Institution:

### 256. A semantic segmentation algorithm for fashion images based on modified

摘要: The semantic segmentation of human body images has huge application potential in many fields, such as autonomous driving, artificial intelligence (AI) face changing, and virtual try-on. Nowadays, many researchers use additional human body posture information to generate multi-level human body analysis images. However, the existing method has limitations when faced with multiple poses and overlapping targets. In this paper, a novel algorithm based on Mask RCNN which has pixel-level accuracy is proposed. In the feature extraction process, a multi-scale feature fusion module applying dilated convolution is proposed to obtain richer semantic information from different perceptual fields. We added a small residual module to the original residual unit structure to increase the size of the receptive field of each layer to capture details and global characteristics. Three convolution kernels with different ratios are designed to obtain receptive fields of different scales. The experimental results show that our method has better performance while considering both object positioning and target classification.  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:000948950300005  
Indexed Date:  
2023-04-05  
Institution Address:  
Jiangnan Univ, Key Lab Ecotext, Minist Educ, 1800 Lihu Ave, Wuxi 214122, Jiangsu, Peoples R China  
Institution:

### 257. A lightweight CNN based information fusion for image denoising

摘要: Deep convolutional neural networks (CNNs) with strong learning abilities have obtained good results for image denoising. However, the CNNs for image denoising have increasingly heavy-weighted, which is not suitable for practical applications. In this paper, we present a lightweight width information fusion CNN(LWIFCNN) for image denoising to address this problem. The proposed model employs two key modules, i.e., multi-scale width information block (MWIB) and information enhancement block (IEB), to improve the model representing capability without heavy complexity. Specifically, in order to extract as much information as possible with low weights, MWIB utilizes a standard convolution and three lightweight residual attention blocks (RABs) to achieve multi-scale feature fusion. Each RAB utilizes two lightweight blocks (LWBs) and an enhanced channel attention mechanism (ECA) to extract width information and reduce computational complexity. IEB uses serial modules and ghost modules to combine width features and depth features to further enhance the representing capability of the model. Experimental results show that our method is better than many excellent methods in both quantitative and qualitative metrics.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001050048600013  
Indexed Date:  
2023-08-30  
Institution Address:  
Harbin Univ Sci & Technol, Sch Econ & Management, Harbin 150006, Heilongjiang, Peoples R China  
Institution Address:  
Cent South Univ, Sch Comp Sci, Changsha 410083, Hunan, Peoples R China  
Institution:  
Institution:

### 258. Real-Time Pantograph Anomaly Detection Using Unsupervised Deep Learning

摘要: An image-based real-time pantograph anomaly detection method is presented by combining unsupervised deep learning and nearest neighbor classification. The proposed method includes the following key steps. First, an improved DeblurGAN-v2 deblurring algorithm is applied to the input pantograph image if there exists motion blur. Next, deep learning semantic segmentation with hybrid coding that combines lightweight convolutional neural network (CNN) and vision transformer (VIT) is employed to accurately segment the pantograph structure within the image. And multiscale feature-dense aggregation network based on an attentional feature fusion (AFF) module has been designed to efficiently integrate information from different feature layers. Finally, a K -nearest neighbor (KNN) classification algorithm with deep pretrained features from the segmented pantograph mask image has been utilized to detect anomalies in the pantograph. Experimental results demonstrate that the proposed pantograph segmentation network outperforms several general segmentation algorithms, achieving a high mean intersection over union (MIoU) of 95.86% with a parameter size of 7 M and FPS of 81.7. And nearest neighbor classification with deep pretrained features achieves excellent pantograph anomaly detection performance with area under the receiver operating characteristic (ROC) curve of 0.987 and area under a precision-recall (PR) curve of 0.998. It is verified that the proposed pantograph anomaly detection method does not rely on abnormal data, and can achieve a high anomaly detection accuracy of 98.75%.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001180920500033  
Indexed Date:  
2024-04-04  
Institution Address:  
Southwest Jiaotong Univ, Sch Elect Engn, Chengdu 611756, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, Sch Phys Sci & Technol, Chengdu 611756, Peoples R China  
Institution Address:  
Zhuzhou CRRC Times Elect Co Ltd, Zhuzhou 412000, Peoples R China  
Institution:  
Institution:  
Institution:

### 259. An Improved Sea Ice Classification Algorithm with Gaofen-3

摘要: The distribution of sea ice is one of the major safety hazards for sea navigation. As human activities in polar regions become more frequent, monitoring and forecasting of sea ice are of great significance. In this paper, we use SAR data from the C-band synthetic aperture radar (SAR) Gaofen-3 satellite in the dual-polarization (VV, VH) fine strip II (FSII) mode of operation to study the Arctic sea ice classification in winter. SAR data we use were taken in the western Arctic Ocean from January to February 2020. We classify the sea ice into four categories, namely new ice (NI), thin first-year ice (tI), thick first-year ice (TI), and old ice (OI), by referring to the ice maps provided by the Canadian Ice Service (CIS). Then, we use the deep learning model MobileNetV3 as the backbone network, input samples of different sizes, and combine the backbone network with multiscale feature fusion methods to build a deep learning model called Multiscale MobileNet (MSMN). Dual-polarization SAR data are used to synthesize pseudocolor images and produce samples of sizes 16 x 16 x 3, 32 x 32 x 3, and 64 x 64 x 3 as input. Ultimately, MSMN can reach over 95% classification accuracy on testing SAR sea ice images. The classification results using only VV polarization or VH polarization data are tested, and it is found that using dual-polarization data could improve the classification accuracy by 10.05% and 9.35%, respectively. When other classification models are trained using the training data from this paper for comparison, the accuracy of MSMN is 4.86% and 1.84% higher on average than that of the model built using convolutional neural networks (CNNs) and ResNet18 model, respectively.  
在中国科学引文数据库中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
24  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
27  
eISSN:  
2072-4292  
入藏号:  
WOS:000771813700001  
Indexed Date:  
2022-03-27  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Chinese Acad Sci, Key Lab Technol Geospatial Informat Proc & Applic, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 101408, Peoples R China  
Institution Address:  
StarWiz Technol Co Ltd, Beijing 100044, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
StarWiz Technol Co Ltd

### 260. CCST: crowd counting with swin transformer

摘要: Accurately estimating the number of individuals contained in an image is the purpose of the crowd counting. It has always faced two major difficulties: uneven distribution of crowd density and large span of head size. Focusing on the former, most CNN-based methods divide the image into multiple patches for processing, ignoring the connection between the patches. For the latter, the multi-scale feature fusion method using feature pyramid ignores the matching relationship between the head size and the hierarchical features. In response to the above issues, we propose a crowd counting network named CCST based on swin transformer, and tailor a feature adaptive fusion regression head called FAFHead. Swin transformer can fully exchange information within and between patches, and effectively alleviate the problem of uneven distribution of crowd density. FAFHead can adaptively fuse multi-level features, improve the matching relationship between head size and feature pyramid hierarchy, and relief the problem of large span of head size available. Experimental results on common datasets show that CCST has better counting performance than all weakly supervised counting works and great majority of popular density map-based fully supervised works.  
在中国科学引文数据库中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
23  
被引频次合计:  
25  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:000785909400001  
Indexed Date:  
2022-04-29  
Institution Address:  
Beijing Univ Technol, Beijing Inst Artificial Intelligence, Dept Informat Sci, Beijing Key Lab Multimedia & Intelligent Software, Beijing 100124, Peoples R China  
Institution Address:  
Beijing Municipal Transportat Operat Coordinat Ct, Beijing 100161, Peoples R China  
Institution:  
Institution:  
Beijing Municipal Transportat Operat Coordinat Ct

### 261. Retinanet\_G2S: a multi-scale feature fusion-based network for fruit

摘要: In the natural environment, the detection and recognition process of Punna navel orange fruit using machine vision systems is affected by many factors, such as complex background, uneven light illumination, occlusions of branches and leaves and large variations in fruit size. To solve these problems of low accuracy in fruit detection and poor robustness of the detection algorithm in the field conditions, a new object detection algorithm, named Retinanet\_G2S, was proposed in this paper based on the modified Retinanet network. The images of Punna navel orange were collected with Microsoft Kinect V2 in the uncontrolled environment. Firstly, a new Res2Net-GF network was designed to replace the section of feature extraction in the original Retinanet, which can potentially improve the learning ability of target features of the trunk network. Secondly, a multi-scale cross-regional feature fusion grids network was designed to replace the feature pyramid network module in the original Retinanet, which could enhance the ability of feature information fusion among different scales of the feature pyramid. Finally, the original border regression localization method in Retinanet network was optimized based on the accurate boundary box regression algorithm. The study results showed that, compared with the original Retinanet network, Retinanet\_G2S improved mAP, mAP50, mAP75, mAPS, mAPM and mAPL by 3.8%, 1.7%, 5.8%, 2.4%, 2.1% and 5.5%, respectively. Moreover, compared with 7 types of classic object detection models, including SSD, YOLOv3, CenterNet, CornerNet, FCOS, Faster-RCNN and Retinanet, the average increase in mAP of Retinanet\_G2S was 9.11%. Overall, Retinanet\_G2S showed a promising optimization effect, particularly for the detection of small targets and overlapping fruits.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
1385-2256  
eISSN:  
1573-1618  
入藏号:  
WOS:001126398800001  
Indexed Date:  
2024-01-01  
Institution Address:  
South China Agr Univ, Coll Math & Informat, Guangzhou 510642, Peoples R China  
Institution Address:  
Foshan Zhongke Innovat Res Inst Intelligent Agr, Foshan 528010, Peoples R China  
Institution Address:  
Mississippi State Univ, Dept Agr & Biol Engn, Mississippi State, MS 39762 USA  
Institution Address:  
South China Univ Technol, Sch Chem & Chem Engn, Guangzhou 510641, Peoples R China  
Institution Address:  
Jiaying Univ, Sch Elect & Informat Engn, Meizhou 514015, Peoples R China  
Institution:  
Institution:  
Foshan Zhongke Innovat Res Inst Intelligent Agr  
Institution:  
Institution:  
Institution:

### 262. Dual-resolution transformer combined with multi-layer separable

摘要: Environmental perception is crucial for unmanned mobile platforms such as autonomous vehicles and robots. Precise and fast semantic segmentation of the surrounding scene is a key task to enhance this capability. Existing real-time semantic segmentation networks are typically based on convolutional neural networks (CNNs), which have achieved good results, but they still lack control over global context features. In recent years, the Transformer architecture has achieved significant success in capturing global context, which is beneficial for improving segmentation accuracy. However, Transformers tend to ignore local connections, and their computational complexity makes real-time segmentation challenging. We propose a lightweight real-time semantic segmentation network called DTMC-Net, which combines the advantages of CNNs and Transformers. We design a special residual convolution module called the Lightweight Multi -layer Separable Convolution Attention module (LMSCA) to reduce the parameter count and perform multi -scale feature fusion to capture local features effectively. We introduce the Simple Dual -Resolution Transformer (SDR Transformer) that utilizes lightweight attention mechanisms and residual feed forward networks to capture and maintain features, with multiple bilateral fusions between two branches to exchange information. The proposed Antiartifact Aggregation Pyramid Pooling Module (AAPPM) optimizes the upsampling process, refines features, and performs multi -scale feature fusion again. DTMC-Net only contains 4.2M parameters and achieves good performance on multiple public datasets with different scenarios.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0097-8493  
eISSN:  
1873-7684  
入藏号:  
WOS:001164120900001  
Indexed Date:  
2024-02-27  
Institution Address:  
Tianjin Univ, Coll Intelligence & Comp, Tianjin 300350, Peoples R China  
Institution:

### 263. Stripe noise removal in high resolution satellite remote sensing images

摘要: Affected by imaging conditions, data transmission, and other factors, stripe noise is common in satellite remote sensing images. It seriously restricts the quality and further use of images. In early studies, various denoising methods, such as statistics-based methods, filtering-based methods, and optimization-based methods, have been proposed to overcome the above problems. These proposed methods have achieved inspiring results in some aspects. However, they still suffer from poor adaptability, low denoising efficiency, and the need for prior knowledge. Therefore, stripe noise removal remains a challenging task. In this study, we take advantage of the convolutional deep network while considering the characteristics of the stripe noise image itself. A deep-learning-based method is proposed, which includes three parts: a feature extraction module, a feature fusion module, and a stripe denoising module. The feature extraction module uses the convolutional layer of the same channel with different strides to extract features. As a result, different-scale feature maps of the noisy image are obtained for the following feature fusion module. The feature fusion module upsamples different-scale feature maps. It fuses these upsampled feature maps through the element-wise addition method. Finally, a denoising network is used to predict the components of stripe noise. The stripe component is subtracted from the noise image based on predictions. Given the difficulties in obtaining real noise samples, the network is trained by simulation samples. Then, it is extended to denoise real images. Experiments on simulation and real images show the excellent performance of our network. In the quantitative assessment, the PSNR and the SSIM of our network when simulated images are used are higher than those of the four methods. In the visual assessment, our network performs well on homogeneous and nonhomogeneous objects. Our network denoises more efficiently and retains more details of ground features than traditional methods and other denoising networks. In real noise images, our method achieves the best denoising performance with the highest ICV and the lowest MRD. Compared with traditional methods, our network has a fast denoising speed, approximately 100 times faster than the denoising speed of the optimization-based denoising method. The above experimental results demonstrate that our network has the best denoising performance in simulated and real images. In this study, a convolutional neural network denoising method based on multiscale feature fusion is proposed based on the fully convolutional neural network. The method uses residual learning to predict the strip-noise components on images. It achieves clean images by subtracting the strip components from noisy images. Experiments demonstrate that compared with traditional methods, deep learning denoising methods are adaptive for removing stripe noise of different intensities without losing image details. The strategy of feature fusion and residual learning can effectively improve the training speed and denoising accuracy of the network. In the future, skip-connected and batch normalization layers will be included to optimize training speed and improve denoising performance. Further studies will be conducted in terms of the transfer ability of the network and the extension of its application in other types of remote sensing images, such as aerial images and hyperspectral images.  
摘要:  
受到成像环境、硬件条件等因素的限制,高分辨率卫星遥感影像上普遍存在条带噪声的现象,其严重影响了影像的辐射质量和可用性。本文针对传统条带去除方法存在的适应性差、去噪效率低、依靠先验知识等不足,提出了一种基于深度学习卷积神经网络的条带噪声去除方法。本方法首先利用不同尺度的卷积层进行特征提取,然后对多尺度的特征图进行特征融合得到去噪底图,通过残差学习的方法在底图上预测存在的噪声分量,最后用噪声影像减去条带噪声分量实现噪声的去除。以模拟和真实获取的噪声影像为实验数据,将本文提出的方法与一些经典的去噪方法进行实验结果对比分析,实验结果表明本文提出的基于深度学习的条带噪声去除方法能够在保留影像地物细节的情况下,能以优异的速度达到最高的定量指标和最好的视觉效果,充分证明了本文方法的优越性。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1007-4619  
入藏号:  
CSCD:7441526  
Indexed Date:  
2023-07-14  
Institution Address:  
武汉大学遥感信息工程学院, 武汉, 湖北 430079, 中国  
Institution Address:  
中国资源卫星应用中心, 北京 100094, 中国  
Institution Address:  
湖北省国土资源研究院, 武汉, 湖北 430071, 中国  
Institution Address:  
School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei 430079, China  
Institution Address:  
China Centre for Resources Satellite Data and Application, Beijing 100094, China  
Institution Address:  
Hubei Provincial Research Institute of Land and Resources, Wuhan, Hubei 430071, China  
Institution:  
武汉大学遥感信息工程学院  
Institution:  
中国资源卫星应用中心  
Institution:  
湖北省国土资源研究院  
Institution:  
School of Remote Sensing and Information Engineering, Wuhan University  
Institution:  
China Centre for Resources Satellite Data and Application  
Institution:  
Hubei Provincial Research Institute of Land and Resources

### 264. Gl-MambaNet: A global-local hybrid Mamba network for medical image

摘要: In recent years, deep learning has made significant progress in medical image segmentation, but it still faces challenges in capturing global features and managing high computational complexity. The Mamba architecture has garnered considerable attention in medical image segmentation tasks due to its ability to capture global information with low computational overhead while maintaining high segmentation accuracy. Based on this, we propose a dual-path network design that combines CNN and Mamba architectures, utilizing a global- local feature fusion module (GLFFM), a multi-stage attention enhancement module (MSAEM), and a semantic feature alignment enhancement module (SFAEM) to significantly improve semantic consistency and multi- scale feature fusion. Specifically, GLFFM is employed to integrate the complementary features of CNN and Mamba, MSAEM enhances feature representation through channel and spatial attention mechanisms, while SFAEM precisely activates relevant features using group attention mechanisms and ensures consistency in dimensions and resolution across skip connection feature maps. Experimental validation on three publicly available medical image datasets, namely GlaS, PH2, and DSB2018, demonstrates that the proposed model, GLMambaNet, outperforms existing state-of-the-art methods in terms of segmentation accuracy, memory usage, and computational efficiency.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:001423915800001  
Indexed Date:  
2025-02-23  
Institution Address:  
Cent South Univ, Sch Comp Sci & Engn, Changsha 410083, Hunan, Peoples R China  
Institution Address:  
Cent South Univ, Big Data Inst, Changsha 410083, Hunan, Peoples R China  
Institution Address:  
Shaoyang Univ, Sch Informat Sci & Engn, Shaoyang 422000, Hunan, Peoples R China  
Institution:  
Institution:  
Institution:

### 265. Combining Multi-Scale Fusion and Attentional Mechanisms for Assessing

摘要: Featured Application This paper focuses on the intelligent application of Shanghai's writing grade examination, aiming to achieve efficient recognition of handwritten Chinese character images in the promotion examination for primary and secondary school students.Abstract Traditional methods of assessing handwritten characters are often too subjective, inefficient, and lagging in feedback, which makes it difficult for educators to achieve fully objective writing assessments and for writers to receive timely suggestions for improvement. In this paper, we propose a convolutional neural network (CNN) architecture that combines the attention mechanism with multi-scale feature fusion; specifically, the features are weighted by designing a bottleneck layer that combines the Squeeze-and-Excitation (SE) attention mechanism to highlight the important information and by applying a multi-scale feature fusion method to enable the network to capture both the global structure and the local details of Chinese characters. Finally, a high-quality dataset containing 26,800 images of handwritten Chinese characters is constructed based on the application scenario of the writing grade test, covering the common Chinese characters in the writing grade exam; The experimental results show that the proposed method achieves 98.6% accuracy on the writing grade exam dataset and 97.05% on the ICDAR-2013 public dataset, significantly improving recognition accuracy. The constructed dataset and improved model are suitable for application scenarios such as writing grade exams, which helps to improve marking efficiency and accuracy.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001418442400001  
Indexed Date:  
2025-02-17  
Institution Address:  
Shanghai Univ Engn Sci, Coll Elect & Elect Engn, Shanghai 201620, Peoples R China  
Institution:

### 266. Image detection of aortic dissection complications based on multi-scale

摘要: Background: Aortic dissection refers to the true and false two-lumen separation of the aortic wall, in which the blood in the aortic lumen enters the aortic mesomembrane from the tear of the aortic intima to separate the mesomembrane and expand along the long axis of the aorta. Purpose: In view of the problems of individual differences, complex complications and many small targets in clinical aortic dissection detection, this paper proposes a convolution neural network MFF-FPN (Multi-scale Feature Fusion based Feature Pyramid Network) for the detection of aortic dissection complications. Methods: The proposed model uses Resnet50 as the backbone for feature extraction and builds a pyramid structure to fuse low-level and high-level feature information. We add an attention mechanism to the backbone network, which can establish inter-dependencies between feature graph channels and enhance the representation quality of CNN. Results: The proposed method has a mean average precision (MAP) of 99.40% in the task of multi object detection for aortic dissection and complications, which is higher than the accuracy of 96.3% on SSD model and 99.05% on YoloV7 model. It greatly improves the accuracy of small target detection such as cysts, making it more suitable for clinical focus detection. Conclusions: The proposed deep learning model achieves feature reuse and focuses on local important information. By adding only a small number of model parameters, we are able to greatly improve the detection accuracy, which is effective in detecting small target lesions commonly found in clinical settings, and also performs well on other medical and natural datasets.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2405-8440  
入藏号:  
WOS:001250952800001  
PubMed ID:  
38533058  
Indexed Date:  
2024-06-25  
Institution Address:  
Cent South Univ Forestry & Technol, Changsha, Hunan, Peoples R China  
Institution Address:  
Cent South Univ, Xiangya Hosp 2, Changsha, Hunan, Peoples R China  
Institution:  
Institution:

### 267. Classification of Microscopic Hyperspectral Images of Blood Cells Based

摘要: Hyperspectral imaging has emerged as a novel imaging modality in the medical field, offering the ability to acquire images of biological tissues while simultaneously providing biochemical insights for in-depth tissue analysis. This approach facilitates early disease diagnosis, presenting advantages over traditional medical imaging techniques. Addressing challenges such as the computational burden of existing convolutional neural networks (CNNs) and imbalances in sample data, this paper introduces a lightweight GhostMRNet for the classification of microscopic hyperspectral images of human blood cells. The proposed model employs Ghost Modules to replace conventional convolutional layers and a cascading approach with small convolutional kernels for multiscale feature extraction, aiming to enhance feature extraction capabilities while reducing computational complexity. Additionally, an SE (Squeeze-and-Excitation) module is introduced to selectively allocate weights to features in each channel, emphasizing informative features and efficiently achieving spatial-spectral feature extraction in microscopic hyperspectral imaging. We evaluated the performance of the proposed GhostMRNet and compared it with other state-of-the-art models using two real medical hyperspectral image datasets. The experimental results demonstrate that GhostMRNet exhibits a superior performance, with an overall accuracy (OA), average accuracy (AA), and Kappa coefficient reaching 99.965%, 99.565%, and 0.9925, respectively. In conclusion, the proposed GhostMRNet achieves a superior classification performance at a smaller computational cost, thereby providing a novel approach for blood cell detection.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2079-9292  
入藏号:  
WOS:001210728300001  
Indexed Date:  
2024-05-10  
Institution Address:  
Hohai Univ, Coll Informat Sci & Engn, Nanjing 211100, Peoples R China  
Institution:

### 268. Intelligent Mining Road Object Detection Based on Multiscale Feature

摘要: In complex mining environments, driverless mining trucks are required to cooperate with multiple intelligent systems. They must perform obstacle avoidance based on factors such as the site road width, obstacle type, vehicle body movement state, and ground concavity-convexity. Targeting the open-pit mining area, this paper proposes an intelligent mining road object detection (IMOD) model developed using a 5G-multi-UAV and a deep learning approach. The IMOD model employs data sensors to monitor surface data in real time within a multisystem collaborative 5G network. The model transmits data to various intelligent systems and edge devices in real time, and the unmanned mining card constructs the driving area on the fly. The IMOD model utilizes a convolutional neural network to identify obstacles in front of driverless mining trucks in real time, optimizing multisystem collaborative control and driverless mining truck scheduling based on obstacle data. Multiple systems cooperate to maneuver around obstacles, including avoiding static obstacles, such as standing and lying dummies, empty oil drums, and vehicles; continuously avoiding multiple obstacles; and avoiding dynamic obstacles such as walking people and moving vehicles. For this study, we independently collected and constructed an obstacle image dataset specific to the mining area, and experimental tests and analyses reveal that the IMOD model maintains a smooth route and stable vehicle movement attitude, ensuring the safety of driverless mining trucks as well as of personnel and equipment in the mining area. The ablation and robustness experiments demonstrate that the IMOD model outperforms the unmodified YOLOv5 model, with an average improvement of approximately 9.4% across multiple performance measures. Additionally, compared with other algorithms, this model shows significant performance improvements.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
eISSN:  
2504-446X  
入藏号:  
WOS:000979326600001  
Indexed Date:  
2023-05-23  
Institution Address:  
China Univ Min & Technol Beijing, Sch Mech Elect & Informat Engn, Beijing 100083, Peoples R China  
Institution Address:  
Beijing Union Univ, Beijing Key Lab Informat Serv Engn, Beijing 100101, Peoples R China  
Institution Address:  
Beijing Inst Technol, Sch Automat, Beijing 100081, Peoples R China  
Institution Address:  
Aerosp Shenzhou Aerial Vehicle Ltd, Tianjin 300301, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Aerosp Shenzhou Aerial Vehicle Ltd

### 269. DATran: Dual Attention Transformer for Multi-Label Image Classification

摘要: Multi-label image classification is a fundamental yet challenging task, which aims to predict the labels associated with a given image. Most of previous methods directly exploit the high-level features from the last layer of convolutional neural network for classification. However, these methods cannot obtain global features due to the limited size of convolutional kernels, and they fail to extract multi-scale features to effectively recognize small-scale objects in the images. Recent studies exploit the graph convolution network to model the label correlations for boosting the classification performance. Despite substantial progress, these methods rely on manually pre-defined graph structures. Besides, they ignore the associations between semantic labels and image regions, and do not fully explore the spatial context of images. To address above issues, we propose a novel Dual Attention Transformer (DATran) model, which adopts a dual-stream architecture that simultaneously learns spatial and channel correlations from multi-label images. Firstly, in order to solve the problem that current methods are difficult to recognize small-size objects, we develop a new multi-scale feature fusion (MSFF) module to generate multi-scale feature representation by jointly integrating both high-level semantics and low-level details. Secondly, we design a prior-enhanced spatial attention (PSA) module to learn the long-range correlation between objects from different spatial positions in images to enhance the model performance. Thirdly, we devise a prior-enhanced channel attention (PCA) module to capture the inter-dependencies between different channel maps, thus effectively improving the correlation between semantic categories. It is worth noting that PSA module and PCA module complement and promote each other to further augment the feature representations. Finally, the outputs of these two attention modules are fused to obtain the final features for classification. Performance evaluation experiments are conducted on MS-COCO 2014, PASCAL VOC 2007 and VG-500 datasets, demonstrating that DATran model achieves better performance than current state-of-the-art models.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:001138814400025  
Indexed Date:  
2024-02-08  
Institution Address:  
Sun Yat Sen Univ, Sch Elect & Informat Technol, Guangzhou 510006, Guangdong, Peoples R China  
Institution:

### 270. Sentinel-1-Based Water and Flood Mapping: Benchmarking Convolutional

摘要: In this study, the effectiveness of several convolutional neural network architectures (AlbuNet-34/FCN/DeepLabV3+/U-Net/U-Net++) for water and flood mapping using Sentinel-1 amplitude data is compared to an operational rule-based processor (S-1FS). This comparison is made using a globally distributed dataset of Sentinel-1 scenes and the corresponding ground truth water masks derived from Sentinel-2 data to evaluate the performance of the classifiers on a global scale in various environmental conditions. The impact of using single versus dual-polarized input data on the segmentation capabilities of AlbuNet-34 is evaluated. The weighted cross entropy loss is combined with the Lovasz loss and various data augmentation methods are investigated. Furthermore, the concept of atrous spatial pyramid pooling used in DeepLabV3+ and the multiscale feature fusion inherent in U-Net++ are assessed. Finally, the generalization capacity of AlbuNet-34 is tested in a realistic flood mapping scenario by using additional data from two flood events and the Sen1Floods11 dataset. The model trained using dual polarized data outperforms the S-1FS significantly and increases the intersection over union (IoU) score by 5%. Using a weighted combination of the cross entropy and the Lovasz loss increases the IoU score by another 2%. Geometric data augmentation degrades the performance while radiometric data augmentation leads to better testing results. FCN/DeepLabV3+/U-Net/U-Net++ perform not significantly different to AlbuNet-34. Models trained on data showing no distinct inundation perform very well in mapping the water extent during two flood events, reaching IoU scores of 0.96 and 0.94, respectively, and perform comparatively well on the Sen1Floods11 dataset.  
在 BIOSIS Citation Index 中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
38  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
38  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000766623900002  
Indexed Date:  
2022-03-24  
Institution Address:  
German Aerosp Ctr DLR, German Remote Sensing Data Ctr DFD, D-82234 Oberpfaffenhofen, Germany  
Institution:

### 271. Multi-scale feature learning via residual dynamic graph convolutional

摘要: Thanks to the relevant spectral-spatial information, hyperspectral images (HSIs) have been widely exploited in Earth observation. Recently, graph convolutional networks (GCNs) have attracted increasing attention in HSI classification due to their advantages in processing non-Euclidean structure data. Unlike convolutional neural networks (CNNs), which perform convolution operations on regular square regions, GCNs can directly work on graph structure data to extract the relationships among adjacent land covers. However, extracting meaningful and deep discriminative spectral-spatial features from HSIs is still a challenging task. In this article, a novel multi-scale feature learning via residual dynamic graph convolutional network is designed for HSI classification, which can extract large-scale contextual spatial structures at superpixel-level graph and local spectral-spatial information at pixel-level, significantly improving the performance of HSI classification. Unlike from the existing GCN-based methods that operate on a graph with a fixed neighbourhood size, multiple graphs with diverse neighbourhood scales are built to comprehensively leverage spectral-spatial information and relationship at multiple scales, and these graphs are dynamically updated to generate more discriminative features (via dynamic GCN) during the convolution process. Moreover, to fully use the multi-scale features extracted from HSIs, a multi-scale feature fusion module is developed to emphasize important features and suppress irrelevant ones. Extensive experiments carried on three benchmark data sets demonstrate the superiority of the proposed approach over other state-of-the-art methods.  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:001151644600001  
Indexed Date:  
2024-02-05  
Institution Address:  
Jiangnan Univ, Sch Artificial Intelligence & Comp Sci, Wuxi, Peoples R China  
Institution Address:  
CNR, Inst Methodol Environm Anal, Potenza, Italy  
Institution Address:  
Univ Grenoble Alpes, GIPSA Lab, Grenoble, France  
Institution:  
Institution:  
Institution:

### 272. MLANet: A Robust Ship Segmentation Network Based on Multilevel

摘要: Synthetic aperture radar (SAR) is a powerful sensor for long-range, all-weather, and large-scale surveillance, making SAR-based ship semantic segmentation a research hotspot. However, accurate segmentation ships, especially small vessels, in near-port waters remain a challenge due to complex background interference in SAR images. Current methods often struggle to extract sufficient features for small ships, leading to high missed detection rates. Furthermore, the complex oceanic background increases false detection rates. To address these issues, we propose MLANet, a multilevel feature-enhanced, multiattention fusion network specifically designed for ship segmentation in SAR images. MLANet leverages the strengths of both convolutional neural network (CNN) and Transformer to perform efficient multiscale feature extraction. The feature enhancement module (FEM) refines global and local features, retaining critical information for small ships, while the attention fusion module reduces background interference. Additionally, a hybrid loss function emphasizes both the shape and boundary of vessels during segmentation. The experimental results show that MLANet achieves 94.83% mean pixel accuracy (mPA) and 90.66% mean intersection over union (mIoU) on the SAR ship detection dataset (SSDD), and 91.74% mPA and 87.72% mIoU on the high-resolution SAR image dataset (HRSID), demonstrating its strong competitiveness and effectiveness in challenging environments.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001397526100046  
Indexed Date:  
2025-02-13  
Institution Address:  
Shandong Univ Sci & Technol, Coll Transportat, Qingdao 266590, Peoples R China  
Institution Address:  
Shandong Univ Sci & Technol, Coll Elect Engn & Automat, Qingdao 266590, Peoples R China  
Institution Address:  
Beijing Jiaotong Univ, Sch Automat & Intelligence, Beijing 100044, Peoples R China  
Institution:  
Institution:  
Institution:

### 273. TransFusion: Efficient Vision Transformer based on 3D transesophageal

摘要: 3D transesophageal echocardiography (TEE) is widely used in the preoperative guidance of left atrial appendage closure (LAAC), and is the preferred imaging examination recommended by expert consensus. The precise extraction of the left atrial appendage (LAA) is the essential initial step for establishing an automated predictive procedure of the LAAC. However, due to the inherent limitations of the ultrasound images and the unique morphological features of the LAA, segmenting the LAA in TEE images is a challenging task. In this paper, we propose a novel Transformer-based dual feature fusion network (TransFusion) for LAA segmentation in TEE images. Our TransFusion includes both intra-stage and inter-stage multi-scale feature fusions. Specifically, in each stage of the encoder, we introduce a Multi-Scale Vision Transformer (MSViT) block for extracting and enhancing multi-scale feature representations. Additionally, we use a Feature Fusion module (FFM) to fuse the features across different stages in parallel and alleviate the semantic gap and better recovering the spatial details. Finally, the convolutional neural network (CNN) as the decoder is used to output the results of LAA segmentation. We compare our method with several state-of-the-art methods on our private clinical dataset. Experimental results demonstrate the superior robustness and segmentation performance of our method and its capability to accurately extract the LAA.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001276428700001  
Indexed Date:  
2024-08-01  
Institution Address:  
Southern Med Univ, Sch Biomed Engn, Guangzhou 510515, Guangdong, Peoples R China  
Institution Address:  
Southern Med Univ, Guangdong Prov Key Lab Med Image Proc, Guangzhou 510515, Guangdong, Peoples R China  
Institution Address:  
Southern Med Univ, Guangdong Prov Engn Lab Med Imaging & Diagnost Tec, Guangzhou 510515, Guangdong, Peoples R China  
Institution Address:  
Southern Med Univ, Zhujiang Hosp, Guangzhou 510280, Peoples R China  
Institution Address:  
Southern Med Univ, Sch Clin Med 2, Guangzhou 510280, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 274. MFBGR: Multi-scale feature boundary graph reasoning network for polyp

摘要: At present, adding Transformer to CNN has promoted the rapid development of colorectal polyp image processing. However, from the perspective of multi-scale feature interaction and boundary coherence, there are mainly some limitations: (1) ignore the local and global correlation within the scale feature, which may cause the missed detection of tiny polyps, (2) lack of multi-scale features to explore the target region, which hinders the learning of multi-variant polyps, and (3) the semantic connection between the target area and the boundary is ignored, cause incoherent segmentation boundaries. In this regard, we design a multi-scale feature boundary graph inference network for polyp segmentation, namely MFBGR. First, the Transformer block captures local- global cues inside the multi-scale information learned by the CNN branches. Second, for the multi-scale global information generated by the Transformer block, we design a cross-scale feature fusion module (CSFM). CSFM performs scale-variation interaction and cascaded fusion to capture the correlation between features across scales and solve the scale-variation problem of segmented objects. Finally, the traditional boundary refinement or enhancement idea is generalized to the graph convolutional reasoning layer (BGRM). BGRM receives CNN's low-level feature information and CSFM's fusion features, or intermediate prediction results, and propagates cross-domain feature information between graph vertices, explores information between target regions and boundary regions, and achieves more accurate boundary segmentation. On the CVC-300, CVC-ClinicDB, CVC-ColonDB, Kvasir-SEG, ETIS datasets, MFBGR and mainstream polyp segmentation networks were compared and tested. MFBGR achieved good results, and Dice, IOU, BAcc, and Haudo were the best. The values reached 94.16%, 89.35% and 97.42%, 3.7442, and the segmentation accuracy of colorectal polyp images has been improved to a certain extent.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
在 BIOSIS Citation Index 中的被引频次:  
3  
被引频次合计:  
11  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:000968693800001  
Indexed Date:  
2023-04-30  
Institution Address:  
Shandong Technol & Business Univ, Coinnovat Ctr Shandong Coll Univ Future Intelligen, Sch Comp Sci & Technol, Yantai 264005, Peoples R China  
Institution Address:  
Shandong Univ Finance & Econ, Sch Comp Sci & Technol, Jinan 250014, Shandong, Peoples R China  
Institution:  
Institution:

### 275. Research on fault diagnosis model of rotating machinery based on

摘要: Existing single-channel networks have poor noise immunity during fault diagnosis of rotating machinery due to the many noises associated with the operation of rotating machinery.To address this problem,a two-channel input LetNet-5 convolutional neural network model incorporating a parallel mechanism was proposed.Case Western Reserve University bearing dataset was used for the model plausibility check process,based on which Gaussian white noise with a signal-to-noise ratio of -10 dB was added to simulate the real noise situation.The shorttime Fourier transform was used to process the motor fan-side and drive-side vibration data,and the resulting timefrequency images were passed to a two-channel input LetNet-5 convolutional neural network for training and learning.The results show that,the dual-channel input LetNet-5 convolutional neural network model is able to capture the fault features in a strong noise environment well,it has higher efficiency and accuracy than the multiscale feature fusion residual model,the multimodal coupled input neural network model,the conventional K-nearest neighbour and decision tree model and the single-channel input LetNet-5 convolutional neural network model.  
摘要:  
针对旋转机械运行过程中伴随着诸多噪声,现有单通道网络在旋转机械故障诊断过程中抗噪性较差的问题,提出了一种加入并联机制的双通道输入LetNet-5卷积神经网络模型。模型合理性检验过程采用了凯斯西储大学轴承数据集,在此基础上,添加信噪比为-10 dB的高斯白噪声模拟真实噪声情形;采用短时傅里叶变换将电机风扇端和驱动端振动数据进行处理,获得的时频图像传递至双通道输入的LetNet-5卷积神经网络进行训练学习。研究结果表明:双通道输入LetNet-5卷积神经网络模型能够良好捕捉到强噪声环境下的故障特征;相比于多尺度特征融合残差模型、多模态耦合输入神经网络模型、传统的K近邻与决策树模型及单通道输入LetNet-5卷积神经网络模型,双通道输入LetNet-5卷积神经网络具有更高的效率和精度。  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1002-3364  
入藏号:  
CSCD:7426060  
Indexed Date:  
2023-06-21  
Institution Address:  
华北电力大学(北京), 电站能量传递转化与系统教育部重点实验室, 北京 102206, 中国  
Institution Address:  
North China Electric Power University(Beijing), Key Laboratory of Power Station Energy Transfer,Transformation And System,Ministry of Education, Beijing 102206, China  
Institution:  
华北电力大学(北京)  
Institution:  
North China Electric Power University(Beijing)

### 276. Lightweight single-image super-resolution via multi-scale feature fusion

摘要: In recent years, single-image super-resolution (SISR) has acquired tremendous progresswith the development of deep learning. However, the majority of SISR methods based on deep learning focus on building more complex networks, which inevitably lead to the problems of computational and memory costs. Thus, these methods may fail to be applied in real-world scenarios. To solve this problem, this paper proposes a lightweight convolution network combined with transformer for SISR named as MMSR. Specifically, an efficient convolutional neural network (CNN) based on multi-scale feature fusion is designed for local feature extraction, which is called MFF-CNN. In addition, we propose a simple and efficient multiple attention block (MAB) to further utilize the context information in features. MAB incorporates channel attention and transformer to help network obtain similar features at a long-term dependence, making full use of global information to further refine texture details. Finally, this paper provides comprehensive results for different settings of the entire network. Experimental results on common used datasets demonstrate that the proposed method can achieve better performances at the 2x, 3x and 4x scales than other state-of-the-art lightweight methods.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
5  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001035766800004  
Indexed Date:  
2023-08-18  
Institution Address:  
Dalian Univ, Sch Software Engn, Natl & Local Joint Engn Lab Comp Aided Design, Dalian 116622, Peoples R China  
Institution Address:  
Dalian Univ Technol, Sch Comp Sci & Technol, Dalian 116024, Peoples R China  
Institution:  
Institution:

### 277. Multi-scale dehazing network via high-frequency feature fusion

摘要: Numerous learning-based methods have achieved significant improvements in haze removal. However, the dehazed results of these methods still suffer from the loss of edge details. To solve this problem, we propose a novel multi-scale dehazing network via high-frequency feature fusion (HFMDN). HFMDN is an end-to-end trainable network, which is mainly composed of four components: a base network (Backbone), a frequency branch network (FBN), a frequency attention module (FAM), and a refine block (RB). The Backbone is a multi-scale feature fusion architecture that can share useful information across different scales. For the training phase, we employ the Laplace Operator to obtain the image's high-frequency (HF) information, which can specifically represent the details of the image (e.g., edges, textures). The FBN takes the HF derived from the original image as an additional prior and utilizes L1 norm loss to constrain the output of FBN to predict the HF of the haze-free image. We further design a frequency attention module (FAM), which automatically learns the weights map of the frequency features to enhance image recovery ability. Furthermore, a refine block (RB) is proposed to extract the features map by fusing the outputs of FBN and Backbone to produce the final haze-free image. The quantitative comparison of the ablation study shows that high-frequency information significantly improves dehazing performance. Extensive experiments also demonstrate that our proposed methods can generate more natural and realistic haze-free images, especially in the contours and details of hazy images. HFMDN performs favorably against the CNN-based state-of-the-art dehazing methods in terms of PSNR, SSIM, and visual effect. (C) 2022 Published by Elsevier Ltd.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
0097-8493  
eISSN:  
1873-7684  
入藏号:  
WOS:000843625800004  
Indexed Date:  
2022-09-10  
Institution Address:  
Guizhou Univ, Coll Comp Sci & Technol, Inst Artificial Intelligence, State Key Lab Publ Big Data, Guiyang 550025, Guizhou, Peoples R China  
Institution Address:  
GuiZhou Educ Univ, Sch Math & Big Data, Guiyang 550001, Guizhou, Peoples R China  
Institution:  
Institution:

### 278. A Deep Learning Model for Small-size Defective Components Detection in

摘要: Unmanned Aerial Vehicle (UAV) inspection has gradually replaced manual inspection of transmission tower, which produces many images. While it is laborious and time-consuming to manually analyze these images, there are also challenges in automatically detecting small-size defective components such as bolts in transmission tower images, due to problems including complex background, small size, and many similar objects of bolts. In this paper, by virtue of multi-scale features and context information, we propose a deep neural network named Camp-Net (Context Information and Multi-Scale Pyramid Network) to identify bolts defect in transmission tower images. First, multi-scale feature fusion combines deep features and shallow features in convolutional networks to detect small-size bolts. Second, context information fusion puts the information around bolts into the detection network to remove the disturbance of complex background and similar objects. An image dataset containing defective bolts and normal bolts is constructed for model training and testing. Experimental results show that bolts with loose pins and bolts without pins among fittings in transmission tower can be accurately identified with the proposed model. The Average Precision (AP) of defective bolts detection of this model can be 11.4% higher than that of the commonly used high performance model, Faster R-CNN.  
Web of Science 核心合集中的 "被引频次":  
18  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
22  
ISSN:  
0885-8977  
eISSN:  
1937-4208  
入藏号:  
WOS:000846890500018  
Indexed Date:  
2022-09-10  
Institution Address:  
North China Elect Power Univ, Sch Control & Comp Engn, Beijing 102206, Peoples R China  
Institution Address:  
Illinois Inst Technol, Dept Elect & Comp Engn, Chicago, IL 60616 USA  
Institution:  
Institution:

### 279. Recognizing Teachers' Hand Gestures for Effective Non-Verbal Interaction

摘要: Hand gesturing is one of the most useful non-verbal behaviors in the classroom, and can help students activate multi-sensory channels to complement teachers' verbal behaviors and ultimately enhance teaching effectiveness. The existing mainstream detection algorithms that can be used to recognize hand gestures suffered from low recognition accuracy under complex backgrounds and different backlight conditions. This study proposes an improved hand gesture recognition framework based on key point statistical transformation features. The proposed framework can effectively reduce the sensitivity of images to background and light conditions. We extracted key points of the image and establish a weak classifier to enhance the anti-interference ability of the algorithm in the case of noise and partial occlusion. Then, we used a deep convolutional neural network model with multi-scale feature fusion to recognize teachers' hand gestures. A series of experiments were conducted on different human gesture datasets to verify the performance of the proposed framework. The results show that the framework proposed in this study has better detection and recognition rates compared to the you only look once (YOLO) algorithm, YOLOv3, and other counterpart algorithms. The proposed framework not only achieved 98.43%, measured by F1 score, for human gesture images in low-light conditions, but also has good robustness in complex lighting environments. We used the proposed framework to recognize teacher gestures in a case classroom setting, and found that the proposed framework outperformed YOLO and YOLOv3 algorithms on small gesture images with respect to recognition performance and robustness.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
2076-3417  
入藏号:  
WOS:000887163900001  
Indexed Date:  
2022-12-06  
Institution Address:  
Quanzhou Normal Univ, High Educ Engn Res Ctr Fujian Prov E Commerce Int, Quanzhou 362000, Peoples R China  
Institution Address:  
Quanzhou Normal Univ, Tan Siu Lin Business Sch, Quanzhou 362000, Peoples R China  
Institution Address:  
Southwest Univ, Fac Educ, Chongqing 400715, Peoples R China  
Institution:  
Institution:  
Institution:

### 280. Jdlmask: joint defogging learning with boundary refinement for foggy

摘要: State-of-the-art instance segmentation approaches, such as Mask R-CNN, have exhibited remarkable performance under clear weather conditions. However, their effectiveness is significantly compromised in foggy environments, primarily due to reduced visibility and obscured object details. To address this challenge, we introduce a joint defogging learning with boundary refinement (JDLMask) framework. Unlike conventional strategies that treat image dehazing as a preprocessing step, JDLMask employs a shared structure that enables the joint learning of defogging and instance segmentation. This integrated approach greatly bolsters the model's adaptability to foggy scenarios. Recognizing challenges in feature extraction due to fog interference, we propose a multi-scale feature fusion mask head, based on the encoder-decoder architecture. This component is designed to acquire both local and global information, thereby enhancing the model's feature representation capacity. Furthermore, we integrate a boundary refinement module, which sharpens the model's localization accuracy by focusing on critical boundary details. Addressing the scarcity of datasets tailored, for instance, segmentation in real-world foggy scenes, we have enriched the Foggy Driving dataset with meticulously crafted instance mask annotations and named it the Foggy Driving InstanceSeg. Comprehensive experiments demonstrate JDLMask's superiority. Compared to the baseline Mask R-CNN, JDLMask achieves improvements of 4.4% and 3.8% in mask AP on the Foggy Cityscapes and Cityscapes validation sets, respectively, and a 2.5% gain on the Foggy Driving InstanceSeg dataset.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001152070300001  
Indexed Date:  
2024-02-06  
Institution Address:  
Tianjin Univ, Sch Elect & Informat Engn, Tianjin 300072, Peoples R China  
Institution Address:  
Shanghai Artificial Intelligence Lab, Shanghai 200232, Peoples R China  
Institution:  
Institution:  
Shanghai Artificial Intelligence Lab

### 281. YOLOv8 Model for Weed Detection in Wheat Fields Based on a Visual

摘要: Accurate weed detection is essential for the precise control of weeds in wheat fields, but weeds and wheat are sheltered from each other, and there is no clear size specification, making it difficult to accurately detect weeds in wheat. To achieve the precise identification of weeds, wheat weed datasets were constructed, and a wheat field weed detection model, YOLOv8-MBM, based on improved YOLOv8s, was proposed. In this study, a lightweight visual converter (MobileViTv3) was introduced into the C2f module to enhance the detection accuracy of the model by integrating input, local (CNN), and global (ViT) features. Secondly, a bidirectional feature pyramid network (BiFPN) was introduced to enhance the performance of multi-scale feature fusion. Furthermore, to address the weak generalization and slow convergence speed of the CIoU loss function for detection tasks, the bounding box regression loss function (MPDIOU) was used instead of the CIoU loss function to improve the convergence speed of the model and further enhance the detection performance. Finally, the model performance was tested on the wheat weed datasets. The experiments show that the YOLOv8-MBM proposed in this paper is superior to Fast R-CNN, YOLOv3, YOLOv4-tiny, YOLOv5s, YOLOv7, YOLOv9, and other mainstream models in regards to detection performance. The accuracy of the improved model reaches 92.7%. Compared with the original YOLOv8s model, the precision, recall, mAP1, and mAP2 are increased by 10.6%, 8.9%, 9.7%, and 9.3%, respectively. In summary, the YOLOv8-MBM model successfully meets the requirements for accurate weed detection in wheat fields.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:001269357000001  
PubMed ID:  
39001158  
Indexed Date:  
2024-07-22  
Institution Address:  
Shandong Agr & Engn Univ, Mech & Elect Engn Coll, Jinan 250100, Peoples R China  
Institution Address:  
Shandong Univ Technol, Sch Agr Engn & Food Sci, Zibo 255000, Peoples R China  
Institution:  
Institution:

### 282. Research on Facial Expression Capture Based on Two-Stage Neural Network

摘要: To generate realistic three-dimensional animation of virtual character, capturing real facial expression is the primary task. Due to diverse facial expressions and complex background, facial landmarks recognized by existing strategies have the problem of deviations and low accuracy. Therefore, a method for facial expression capture based on two-stage neural network is proposed in this paper which takes advantage of improved multi-task cascaded convolutional networks (MTCNN) and high-resolution network. Firstly, the convolution operation of traditional MTCNN is improved. The face information in the input image is quickly filtered by feature fusion in the first stage and Octave Convolution instead of the original ones is introduced into in the second stage to enhance the feature extraction ability of the network, which further rejects a large number of false candidates. The model outputs more accurate facial candidate windows for better landmarks recognition and locates the faces. Then the images cropped after face detection are input into high-resolution network. Multi-scale feature fusion is realized by parallel connection of multi-resolution streams, and rich high-resolution heatmaps of facial landmarks are obtained. Finally, the changes of facial landmarks recognized are tracked in real-time. The expression parameters are extracted and transmitted to Unity3D engine to drive the virtual character's face, which can realize facial expression synchronous animation. Extensive experimental results obtained on the WFLW database demonstrate the superiority of the proposed method in terms of accuracy and robustness, especially for diverse expressions and complex background. The method can accurately capture facial expression and generate three-dimensional animation effects, making online entertainment and social interaction more immersive in shared virtual space.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:000792398000030  
Indexed Date:  
2022-06-02  
Institution Address:  
Hebei Univ Sci & Technol, Sch Informat Sci & Engn, Shijiazhuang 050000, Hebei, Peoples R China  
Institution Address:  
Newcastle Univ, Sch Engn, Newcastle Upon Tyne NE98, Tyne & Wear, England  
Institution:  
Institution:

### 283. End-to-End Ship Detection in SAR Images for Complex Scenes Based on Deep

摘要: Ship detection on synthetic aperture radar (SAR) imagery has many valuable applications for both civil and military fields and has received extraordinary attention in recent years. The traditional detection methods are insensitive to multiscale ships and usually time-consuming, results in low detection accuracy and limitation for real-time processing. To balance the accuracy and speed, an end-to-end ship detection method for complex inshore and offshore scenes based on deep convolutional neural networks (CNNs) is proposed in this paper. First, the SAR images are divided into different grids, and the anchor boxes are predefined based on the responsible grids for dense ship prediction. Then, Darknet-53 with residual units is adopted as a backbone to extract features, and a top-down pyramid structure is added for multiscale feature fusion with concatenation. By this means, abundant hierarchical features containing both spatial and semantic information are extracted. Meanwhile, the strategies such as soft nonmaximum suppression (Soft-NMS), mix-up and mosaic data augmentation, multiscale training, and hybrid optimization are used for performance enhancement. Besides, the model is trained from scratch to avoid learning objective bias of pretraining. The proposed one-stage method adopts end-to-end inference by a single network, so the detection speed can be guaranteed due to the concise paradigm. Extensive experiments are performed on the public SAR ship detection dataset (SSDD), and the results show that the method can detect both inshore and offshore ships with higher accuracy than other mainstream methods, yielding the accuracy with an average of 95.52%, and the detection speed is quite fast with about 72 frames per second (FPS). The actual Sentinel-1 and Gaofen-3 data are utilized for verification, and the detection results also show the effectiveness and robustness of the method.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
18  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
18  
ISSN:  
1687-725X  
eISSN:  
1687-7268  
入藏号:  
WOS:000637362600002  
Indexed Date:  
2021-04-28  
Institution Address:  
Chinese Acad Sci, Inst Microelect, Beijing 100029, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100029, Peoples R China  
Institution:  
Institution:

### 284. Local-Global Based High-Resolution Spatial-Spectral Representation

摘要: Due to the inability of convolutional neural networks to effectively obtain long-range information, a transformer was recently introduced into the field of pansharpening to obtain global dependencies. However, a transformer does not pay enough attention to the information of channel dimensions. To solve this problem, a local-global-based high-resolution spatial-spectral representation network (LG-HSSRN) is proposed to fully fuse local and global spatial-spectral information at different scales. In this paper, a multi-scale feature fusion (MSFF) architecture is designed to obtain the scale information of remote sensing images. Meanwhile, in order to learn spatial texture information and spectral information effectively, a local-global feature extraction (LGFE) module is proposed to capture the local and global dependencies in the source images from a spatial-spectral perspective. In addition, a multi-scale contextual aggregation (MSCA) module is proposed to weave hierarchical information with high representational power. The results of three satellite datasets show that the proposed method exhibits superior performance in terms of both spatial and spectral preservation compared to other methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:000843287100001  
Indexed Date:  
2022-08-28  
Institution Address:  
Zhengzhou Univ Light Ind, Coll Comp & Commun Engn, Zhengzhou 450002, Peoples R China  
Institution:

### 285. MBT-UNet: Multi-Branch Transform Combined with UNet for Semantic

摘要: Remote sensing (RS) images play an indispensable role in many key fields such as environmental monitoring, precision agriculture, and urban resource management. Traditional deep convolutional neural networks have the problem of limited receptive fields. To address this problem, this paper introduces a hybrid network model that combines the advantages of CNN and Transformer, called MBT-UNet. First, a multi-branch encoder design based on the pyramid vision transformer (PVT) is proposed to effectively capture multi-scale feature information; second, an efficient feature fusion module (FFM) is proposed to optimize the collaboration and integration of features at different scales; finally, in the decoder stage, a multi-scale upsampling module (MSUM) is proposed to further refine the segmentation results and enhance segmentation accuracy. We conduct experiments on the ISPRS Vaihingen dataset, the Potsdam dataset, the LoveDA dataset, and the UAVid dataset. Experimental results show that MBT-UNet surpasses state-of-the-art algorithms in key performance indicators, confirming its superior performance in high-precision remote sensing image segmentation tasks.  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
2072-4292  
入藏号:  
WOS:001286931000001  
Indexed Date:  
2024-08-15  
Institution Address:  
Harbin Engn Univ, Coll Intelligent Syst Sci & Engn, Harbin 150001, Peoples R China  
Institution Address:  
Univ Western Australia, Sch Elect Elect & Comp Engn, Perth 6009, Australia  
Institution:  
Institution:

### 286. STMSF: Swin Transformer with Multi-Scale Fusion for Remote Sensing Scene

摘要: Emerging vision transformers (ViTs) are more powerful in modeling long-range dependences of features than conventional deep convolution neural networks (CNNs). Thus, they outperform CNNs in several computer vision tasks. However, existing ViTs fail to encounter the multi-scale characteristics of ground objects with various spatial sizes when they are applied to remote sensing (RS) scene images. Therefore, in this paper, a Swin transformer with multi-scale fusion (STMSF) is proposed to alleviate such an issue. Specifically, a multi-scale feature fusion module is proposed, so that features of ground objects at different scales in the RS scene can be well considered by merging multi-scale features. Moreover, a spatial attention pyramid network (SAPN) is designed to enhance the context of coarse features extracted with the transformer and further improve the network's representation ability of multi-scale features. Experimental results over three benchmark RS scene datasets demonstrate that the proposed network obviously outperforms several state-of-the-art CNN-based and transformer-based approaches.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001430805000001  
Indexed Date:  
2025-03-03  
Institution Address:  
Hanshan Normal Univ, Sch Math & Stat, Chaozhou 521041, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Sch Elect & Informat, Xian 710072, Peoples R China  
Institution:  
Institution:

### 287. Semantic segmentation model for concrete cracks based on parallel

摘要: In recent years, crack detection has been the focus of relevant research since concrete fractures are the most dangerous damage to structures. Computer vision-based approaches are frequently employed for their distinct benefits. However, the crack segmentation model based only on convolutional neural networks (CNNs) is still inadequate in generalization because of its inherent bias produced by its low contextual understanding capacity. This research provides a framework PSC (Parallel Swin-CNNs) that employs a multi-scale feature fusion pyramid decoder to partition concrete cracks semantically using Swin Transformer and CNNs. This research evaluates classic CNN models U-Net, U-Net++, DeepLabV3, PSPNet, Feature Pyramid Network (FPN), and DeepCrack on two datasets. The proposed PSC model achieves the best crack segmentation results, with a maximum improvement of 36.57% in F1-score and 62.38% in Intersection over Union value on both datasets, a reduction in parameters of 2.95%-40.89% except for PSPNet. The proposed PSC model demonstrates versatile applicability across various scenarios, effectively overcoming interferences such as light shadows, oil stains, potholes, and textured surfaces while maintaining high computational efficiency.  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1475-9217  
eISSN:  
1741-3168  
入藏号:  
WOS:001163905500001  
Indexed Date:  
2024-02-28  
Institution Address:  
Nanjing Tech Univ, Coll Civil Engn, Nanjing, Jiangsu, Peoples R China  
Institution Address:  
Yangzhou Univ, Coll Architecture Sci & Engn, Yangzhou 225127, Jiangsu, Peoples R China  
Institution Address:  
Nanjing Univ Technol Chuzhou Co Ltd, Transportat Sci Inst, Chuzhou, Anhui, Peoples R China  
Institution Address:  
Univ Arizona, Dept Phys, Tucson, AZ USA  
Institution Address:  
Southeast Univ, Sch Transportat, Nanjing, Jiangsu, Peoples R China  
Institution:  
Institution:  
Institution:  
Nanjing Univ Technol Chuzhou Co Ltd  
Institution:  
Institution:

### 288. A novel convolutional neural network with interference suppression for

摘要: Despite some recent achievements in the intelligent data-driven fault diagnosis of mechanical rotating components, in most cases very large amounts of pure data are demanded for model training. However, the collected vibration signals of mechanical rotating components are inevitably contaminated with noise in real industries. To resolve the low fault diagnosis accuracy due to strong noise interference, a convolutional neural network with interference suppression (ISCNN) is proposed in this paper. First, a parallel convolutional structure with dilated wide convolution kernels is designed to extract long-time correlated fault features of multiple scales from a noise-contaminated signal by sparse sampling, and noise interference is suppressed by filtering operations. Then, a multiscale feature enhancement module is constructed to achieve adaptive tuning of time scale by employing a parallel selective kernel network to exploit weak fault features buried in noisy signals. Finally, a convolutional feature fusion method is adopted to integrate multidimensional fault features and feed them into the classifier, thus achieving accurate fault diagnosis of mechanical rotating components. Experimental results on two benchmark datasets indicate that the ISCNN model outperforms its counterparts in the fault diagnosis field and improves the fault discriminability of mechanical rotating components in strong noisy scenarios.  
Web of Science 核心合集中的 "被引频次":  
10  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
10  
ISSN:  
0941-0643  
eISSN:  
1433-3058  
入藏号:  
WOS:000759356100001  
Indexed Date:  
2022-03-15  
Institution Address:  
Harbin Inst Technol, Automat Test & Control Inst, Harbin 150080, Peoples R China  
Institution:

### 289. An Improved U-Net Infrared Small Target Detection Algorithm Based on

摘要: Infrared small target detection technology plays a crucial role in various fields such as military reconnaissance, power patrol, medical diagnosis, and security. The advancement of deep learning has led to the success of convolutional neural networks in target segmentation. However, due to challenges like small target scales, weak signals, and strong background interference in infrared images, convolutional neural networks often face issues like leakage and misdetection in small target segmentation tasks. To address this, an enhanced U-Net method called MST-UNet is proposed, the method combines multi-scale feature decomposition and fusion and attention mechanisms. The method involves using Haar wavelet transform instead of maximum pooling for downsampling in the encoder to minimize feature loss and enhance feature utilization. Additionally, a multi-scale residual unit is introduced to extract contextual information at different scales, improving sensory field and feature expression. The inclusion of a triple attention mechanism in the encoder structure further enhances multidimensional information utilization and feature recovery by the decoder. Experimental analysis on the NUDT-SIRST dataset demonstrates that the proposed method significantly improves target contour accuracy and segmentation precision, achieving IoU and nIoU values of 80.09% and 80.19%, respectively.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
eISSN:  
1424-8220  
入藏号:  
WOS:001269370100001  
PubMed ID:  
39001006  
Indexed Date:  
2024-07-22  
Institution Address:  
Guangxi Univ Sci & Technol, Sch Automat, Liuzhou 545006, Peoples R China  
Institution Address:  
Guangxi Univ Sci & Technol, Guangxi Collaborat Innovat Ctr Earthmoving Machine, Liuzhou 545006, Peoples R China  
Institution:  
Institution:

### 290. Multi-load Short-term Joint Forecasting of Integrated Energy System

摘要: A multi-load short-term joint forecasting model of integrated energy system is proposed to improve the accuracy of multi-load forecasting of integrated energy system based on multi-scale feature extraction by comprehensively considering the interaction mechanism of multi-energy, the coupling characteristics of multi-load and the correlation of meteorological factors. Firstly, the coupling characteristics of multi-load and the correlation of influencing factors are studied by the maximum information coefficient, and the prediction characteristics are selected. Secondly, the input features are decomposed by variational modal decomposition technology to enhance features purity. Finally, the CNN-BiLSTM multi-task learning model is used for feature fusion, and the Attention mechanism is used to select important features differently to realize multi-scale feature extraction. In addition, hyperparameter optimization of the VMD and CNN-BiLSTM multi-task learning model is achieved by the snow ablation optimizer to realize joint forecasting of IES multivariate loads. Experiments were conducted using real-world data from Arizona, USA. The results indicate that the proposed joint forecasting method possesses lower root mean square error and higher accuracy compared with single forecasting method or other models and greater robustness in IES multivariate load forecasting.  
摘要:  
为提高综合能源系统(integrated energy system, IES)多元负荷预测的精确度,综合考虑多能源相互作用机理、多元负荷耦合特性及气象因素相关性,提出了一种基于多尺度特征提取的IES多元负荷短期联合预测方法。首先,通过最大互信息系数(maximum information coefficient,MIC)研究多元负荷耦合特性及影响因素相关性,选择预测特征;其次,利用变分模态分解技术(variational mode decomposition, VMD)对输入特征进行分解,提升特征纯洁度;最后,采用卷积神经网络-双向长短期记忆神经网络(convolutional neural network-bidirectional long and short-term memory, CNN-BiLSTM)多任务学习模型进行纵向、横向特征选择,注意力(Attention)机制对重要特征差异化提取,实现多尺度特征提取,并利用雪消融优化器(snow ablation optmizer, SAO)对VMD和CNN-BiLSTM多任务学习模型进行超参数优化,以此实现IES多元负荷的联合预测。以美国亚利桑那州实测数据进行实验,结果表明,无论与单一预测方法还是与其他模型相比,所提联合预测方法的均方根误差更低、准确率更高,在IES多元负荷预测中具有更高的精确性和鲁棒性。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1003-6520  
入藏号:  
CSCD:7769330  
Indexed Date:  
2024-09-27  
Institution Address:  
昆明理工大学电力工程学院, 昆明, 云南 650500, 中国  
Institution Address:  
上海交通大学电子信息与电气工程学院, 上海 200240, 中国  
Institution Address:  
云南电网有限责任公司电力科学研究院, 昆明, 云南 650521, 中国  
Institution Address:  
Faculty of Electric Power Engineering, Kunming University of Science and Technology, Kunming, Yunnan 650500, China  
Institution Address:  
School of Electric Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China  
Institution Address:  
Electric Power Research Institute, Yunnan Power Grid Co., Ltd., Kunming, Yunnan 650521, China  
Institution:  
昆明理工大学电力工程学院  
Institution:  
上海交通大学电子信息与电气工程学院  
Institution:  
云南电网有限责任公司电力科学研究院  
Institution:  
Faculty of Electric Power Engineering, Kunming University of Science and Technology  
Institution:  
School of Electric Information and Electrical Engineering, Shanghai Jiao Tong University  
Institution:  
Electric Power Research Institute, Yunnan Power Grid Co., Ltd.

### 291. Multi-Label Auroral Image Classification Based on CNN and Transformer.

摘要: Auroral image classification has long been a focus of research in auroral physics. However, current methods for automatic auroral classification typically assume that only one type of aurora is present in an auroral image. This oversight neglects the complex transition states and coexistence of multiple types during the auroral evolution process, thus limiting the exploration of the intricate semantics of auroral images. To fully exploit the physical information embedded in auroral images, this paper proposes a multi-label auroral classification method, termed MLAC, which integrates convolutional neural network (CNN) and Transformer architectures. Firstly, we introduce a multi-scale feature fusion framework that enables the model to capture both fine-grained features and high-level information in auroral images, resulting in a more comprehensive representation of auroral features. Secondly, we propose a lightweight multi-head self-attention mechanism that captures long-range dependencies between pixels during the multiscale feature fusion process, which is crucial for distinguishing subtle differences between auroral types. Furthermore, we design a residual focused multilayer perceptron module that integrates large kernel depth-wise convolution with an improved multilayer perceptron. This integration enhances the model's ability to represent complex spatial structure, thus improving local feature extraction and global contextual understanding. The proposed method achieves a mean average precision (mAP) of 88.20% on the auroral observation data collected at the Yellow River Station from 2003 to 2008. This performance significantly surpasses that of the most advanced multi-label classification models while maintaining competitive computational efficiency. Moreover, our method also outperforms the state-of-the-art multi-label methods in both computational efficiency and classification accuracy on two publicly available multi-label image datasets: WIDER-Attribute and VOC2007. These results demonstrate that our method skillfully leverages the robust feature extraction capability of CNNs for local features and the superior global information processing capability of Transformer.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1941-0042  
入藏号:  
MEDLINE:40095833  
PubMed ID:  
40095833  
Indexed Date:  
2025-03-21

### 292. A Collaborative Neural Network-Based Automatic Object Recognition Method

摘要: The widespread application of virtual reality has made automatic target recognition in environmental space particularly important. However, traditional image recognition methods faced challenges due to the complex and diverse environmental characteristics in virtual reality scenes. It is easily influenced by occlusion, noise, and other factors. To deal with the issue, this paper utilizes the combination of convolutional neural networks (CNNs) and recurrent neural networks (RNNs), and proposes a collaborative neural network-based automatic object recognition method for environmental space under virtual reality scenes. Specifically, CNN is mainly responsible for feature extraction of environmental scene images, while RNN is used to model the spatial relationships of image sequences. Then, a data augmentation method is introduced in virtual reality scenarios, which can generate more training samples on a limited dataset, thereby improving the model's generalization ability. Meanwhile, in order to address issues such as occlusion and noise, a multi-scale feature fusion strategy is further developed to improve the robustness of target recognition by multiscale feature fusion. Through experimental verification using actual data collected in virtual reality scenes, the results show that the proposed method has achieved excellent performance and strong adaptability compared with traditional methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0218-1266  
eISSN:  
1793-6454  
入藏号:  
WOS:001361579000002  
Indexed Date:  
2024-11-30  
Institution Address:  
Coll Civil Engn, Zhengzhou Coll Finance & Econ, Zhengzhou 450000, Peoples R China  
Institution Address:  
King Saud Univ, Community Coll, Comp Sci Dept, Riyadh 11437, Saudi Arabia  
Institution:  
Coll Civil Engn  
Institution:

### 293. Lightweight Vehicle Detection Based on Mamba\_ViT

摘要: Vehicle detection algorithms are essential for intelligent traffic management and autonomous driving systems. Current vehicle detection algorithms largely rely on deep learning techniques, enabling the automatic extraction of vehicle image features through convolutional neural networks (CNNs). However, in real traffic scenarios, relying only on a single feature extraction unit makes it difficult to fully understand the vehicle information in the traffic scenario, thus affecting the vehicle detection effect. To address this issue, we propose a lightweight vehicle detection algorithm based on Mamba\_ViT. First, we introduce a new feature extraction architecture (Mamba\_ViT) that separates shallow and deep features and processes them independently to obtain a more complete contextual representation, ensuring comprehensive and accurate feature extraction. Additionally, a multi-scale feature fusion mechanism is employed to enhance the integration of shallow and deep features, leading to the development of a vehicle detection algorithm named Mamba\_ViT\_YOLO. The experimental results on the UA-DETRAC dataset show that our proposed algorithm improves mAP@50 by 3.2% compared to the latest YOLOv8 algorithm, while using only 60% of the model parameters.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001366069300001  
PubMed ID:  
39598916  
Indexed Date:  
2024-12-05  
Institution Address:  
Shandong Jiaotong Univ, Sch Informat & Elect Engn, Jinan 250357, Peoples R China  
Institution:

### 294. Balanced-YOLOv3: Addressing the Imbalance Problem of Object Detection in

摘要: The object detection algorithm of the PCB (Printed Circuit Board) assembly scene based on CNN (Convolutional Neural Network) can significantly improve the production capacity of intelligent manufacturing of electronic products. However, the object class imbalance in the PCB assembly scene, the multi-scale feature imbalance, and the positive/negative sample imbalance in the CNN have become critical problems restricting object detection performance. Based on YOLOv3, this paper proposes a class-balanced Train/Val (Training set/Validation set) split method for object class imbalance, an additional feature fusion strategy for multi-scale feature imbalance, and an efficient anchor concept for positive/negative sample imbalance. These three contributions are Balanced-YOLOv3. After experimental verification, compared with other YOLOv3 series algorithms, the mAP@.5 (Mean Average Precision at Intersection over Union threshold 0.5) and mAP@.5:.95 (average mAP over different Intersection over Union thresholds, from 0.5 to 0.95, step 0.05) of Balanced-YOLOv3 have achieved the best results and ranked third in the metrics of parameter and inference time. Compared with other current anchor-based object detection algorithms, Balanced-YOLOv3 has excellent detection performance and low computational complexity, which effectively solves the problem of imbalanced object detection in PCB assembly scenarios.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
eISSN:  
2079-9292  
入藏号:  
WOS:000787042700001  
Indexed Date:  
2022-05-04  
Institution Address:  
Jiangsu Univ, Sch Mech Engn, Zhenjiang 212000, Jiangsu, Peoples R China  
Institution Address:  
Anyang Inst Technol, Sch Elect Informat & Elect Engn, Anyang 455000, Peoples R China  
Institution Address:  
Guangdong Univ Technol, Sch Mech & Elect Engn, Guangzhou 510006, Peoples R China  
Institution:  
Institution:  
Institution:

### 295. TMNet: A Two-Branch Multi-Scale Semantic Segmentation Network for Remote

摘要: Pixel-level information of remote sensing images is of great value in many fields. CNN has a strong ability to extract image backbone features, but due to the localization of convolution operation, it is challenging to directly obtain global feature information and contextual semantic interaction, which makes it difficult for a pure CNN model to obtain higher precision results in semantic segmentation of remote sensing images. Inspired by the Swin Transformer with global feature coding capability, we design a two-branch multi-scale semantic segmentation network (TMNet) for remote sensing images. The network adopts the structure of a double encoder and a decoder. The Swin Transformer is used to increase the ability to extract global feature information. A multi-scale feature fusion module (MFM) is designed to merge shallow spatial features from images of different scales into deep features. In addition, the feature enhancement module (FEM) and channel enhancement module (CEM) are proposed and added to the dual encoder to enhance the feature extraction. Experiments were conducted on the WHDLD and Potsdam datasets to verify the excellent performance of TMNet.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
4  
eISSN:  
1424-8220  
入藏号:  
WOS:001028610300001  
PubMed ID:  
37447759  
Indexed Date:  
2023-08-06  
Institution Address:  
Inner Mongolia Agr Univ, Sch Comp & Informat Engn, Hohhot 010011, Peoples R China  
Institution Address:  
Inner Mongolia Autonomous Reg Key Lab Big Data Res, Hohhot 750306, Peoples R China  
Institution Address:  
Inner Mongolia Agr Univ, Coll Water Conservancy & Civil Engn, Hohhot 010018, Peoples R China  
Institution Address:  
Key Lab Water Resources Protect & Utilizat Inner M, Hohhot 750306, Peoples R China  
Institution Address:  
Inst Grassland Res CAAS, Hohhot 010013, Peoples R China  
Institution:  
Institution:  
Inner Mongolia Autonomous Reg Key Lab Big Data Res  
Institution:  
Institution:  
Key Lab Water Resources Protect & Utilizat Inner M  
Institution:

### 296. Enhancing Video Surveillance and Behavior Recognition With Deep Learning

摘要: With the rapid advancement of technology, deep learning has demonstrated significant application potential in the fields of video surveillance and behavior recognition. This paper explores the application of an improved Convolutional Neural Network (CNN)-based deep learning model in video behavior recognition and discusses privacy protection strategies. By incorporating multi-scale feature fusion mechanisms, spatiotemporal attention mechanisms, and Long Short-Term Memory (LSTM) networks, the designed deep learning model achieved notable performance improvements on the UCF-101 dataset, with an accuracy rate of 95.8%, precision of 96.5%, recall of 95.2%, and an F1 score of 95.8%. Additionally, this paper outlines a comprehensive privacy protection strategy that includes data anonymization, encrypted transmission, and access control, effectively safeguarding personal privacy. Experimental results indicate that although the privacy protection measures led to an increase in data processing time by approximately 40% and a decrease in model performance by about 2.5%, they significantly enhanced data security. This study provides valuable insights and references for the application of deep learning in video surveillance and behavior recognition, as well as its associated privacy protection.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001346625400001  
Indexed Date:  
2024-11-08  
Institution Address:  
Guangzhou Songtian Polytech Coll, Guangzhou 511370, Guangdong, Peoples R China  
Institution:  
Guangzhou Songtian Polytech Coll

### 297. Single image super-resolution via deep progressive multi-scale fusion

摘要: Deep convolutional neural network-based single-image super-resolution (SR) models typically process either upsampled full-resolution or original low-resolution features, which suffer from context lack and spatially imprecision, respectively. To solve this, we propose a novel progressive SR network to preserve spatial precision through the original resolution and to receive rich contextual information from low-to-high resolution representations. Our proposed progressive, selective scale fusion network includes four key points: (a) parallel multi-scale convolution branches to extract multi-scale features, (b) information exchange across the multi-scale branches, (c) attention mechanism-based multi-scale feature fusion, and (d) gradual aggregation of multi-scale streams from low-to-high resolutions. The proposed method learns hierarchical features that aggregate contextual information from different resolution streams while maintaining high-resolution spatial details. Both quantitative and qualitative experiments on benchmark and real-world datasets show that our method offers a favorable performance against state-of-the-art methods for SR tasks with different scaling factors.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
0941-0643  
eISSN:  
1433-3058  
入藏号:  
WOS:000783461800002  
Indexed Date:  
2022-04-26  
Institution Address:  
East China Jiaotong Univ, Sch Informat Engn, Nanchang 330013, Jiangxi, Peoples R China  
Institution Address:  
Jeonbuk Natl Univ, Div Comp Sci & Engn, CAIIT, Jeonju 54896, South Korea  
Institution:  
Institution:

### 298. Underwater Biological Detection Algorithm Based on Improved Faster-RCNN

摘要: Underwater organisms are an important part of the underwater ecological environment. More and more attention has been paid to the perception of underwater ecological environment by intelligent means, such as machine vision. However, many objective reasons affect the accuracy of underwater biological detection, such as the low-quality image, different sizes or shapes, and overlapping or occlusion of underwater organisms. Therefore, this paper proposes an underwater biological detection algorithm based on improved Faster-RCNN. Firstly, the ResNet is used as the backbone feature extraction network of Faster-RCNN. Then, BiFPN (Bidirectional Feature Pyramid Network) is used to build a ResNet-BiFPN structure which can improve the capability of feature extraction and multi-scale feature fusion. Additionally, EIoU (Effective IoU) is used to replace IoU to reduce the proportion of redundant bounding boxes in the training data. Moreover, K-means++ clustering is used to generate more suitable anchor boxes to improve detection accuracy. Finally, the experimental results show that the detection accuracy of underwater biological detection algorithm based on improved Faster-RCNN on URPC2018 dataset is improved to 88.94%, which is 8.26% higher than Faster-RCNN. The results fully prove the effectiveness of the proposed algorithm.  
在中国科学引文数据库中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
22  
在 BIOSIS Citation Index 中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
26  
eISSN:  
2073-4441  
入藏号:  
WOS:000694471500001  
Indexed Date:  
2021-09-28  
Institution Address:  
Hohai Univ, Jiangsu Key Lab Power Transmiss & Distribut Equip, Changzhou 213022, Peoples R China  
Institution Address:  
Hohai Univ, Coll Internet Things Engn, Changzhou 213022, Peoples R China  
Institution:  
Institution:

### 299. NSD-SSD: A Novel Real-Time Ship Detector Based on Convolutional Neural

摘要: With the rapid development of the marine industry, intelligent ship detection plays a very important role in the marine traffic safety and the port management. Current detection methods mainly focus on synthetic aperture radar (SAR) images, which is of great significance to the field of ship detection. However, these methods sometimes cannot meet the real-time requirement. To solve the problems, a novel ship detection network based on SSD (Single Shot Detector), named NSD-SSD, is proposed in this paper. Nowadays, the surveillance system is widely used in the indoor and outdoor environment, and its combination with deep learning greatly promotes the development of intelligent object detection and recognition. The NSD-SSD uses visual images captured by surveillance cameras to achieve real-time detection and further improves detection performance. First, dilated convolution and multiscale feature fusion are combined to improve the small objects' performance and detection accuracy. Second, an improved prediction module is introduced to enhance deeper feature extraction ability of the model, and the mean Average Precision (mAP) and recall are significant improved. Finally, the prior boxes are reconstructed by using the K-means clustering algorithm, the Intersection-over-Union (IoU) is higher, and the visual effect is better. The experimental results based on ship images show that the mAP and recall can reach 89.3% and 93.6%, respectively, which outperforms the representative model (Faster R-CNN, SSD, and YOLOv3). Moreover, our model's FPS is 45, which can meet real-time detection acquirement well. Hence, the proposed method has the better overall performance and achieves higher detection efficiency and better robustness.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
1687-5265  
eISSN:  
1687-5273  
入藏号:  
WOS:000700351800001  
PubMed ID:  
34557225  
Indexed Date:  
2021-09-14  
Institution Address:  
Shanghai Maritime Univ, Coll Informat Engn, Shanghai 201306, Peoples R China  
Institution Address:  
Shandong Normal Univ, Sch Phys & Elect, Jinan 250358, Peoples R China  
Institution:  
Institution:

### 300. A Lightweight Convolutional Neural Network for Ship Target Detection in

摘要: Synthetic aperture radar (SAR), due to its merits of all-day, all-weather, and high resolution, has become an important component of radar research. Target detection is a significant basis for radar image interpretation. Aiming at the problems, such as low accuracy and high complexity of traditional convolutional neural network in SAR ship target detection, which is not conducive to the deployment of terminal equipment, we proposed a lightweight network for SAR image ship target detection based on the latest anchor-free algorithm YOLOX. First, we redesign a novel lightweight block with stronger feature fusion ability, namely, MobileNetV3S. On the basis of MobileNetV3S, we combine Cross Stage Partial Network to construct a lightweight backbone. Second, in order to improve the ability of multiscale feature extraction, we propose a new module based on dilated convolution with different dilated rates and Efficient Spatial Pyramid Network. Furthermore, we adopt convolutional block attentional module to optimize traditional YOLOX's Feature Pyramid Network, and propose a lightweight enhanced feature extraction module, which can improve the focusing ability of important targets. For detection head, the depth-separable convolution is also applied to reduce the network's parameters. Finally, in terms of loss function, we abandon the traditional Intersection over Union and use absolute Intersection over Union with the better convergence effect. The experimental results on the SAR Ship Detection Dataset show that compared with the baseline YOLOX, although parameters of our method are decreased by 66.7%, its AP reaches 90.8%, which exceeds the baseline YOLOX by 0.5%, and its false detection rate is also obviously reduced, achieving state-of-the-art performance.  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
0018-9251  
eISSN:  
1557-9603  
入藏号:  
WOS:001233499100053  
Indexed Date:  
2024-06-13  
Institution Address:  
Shanghai Maritime Univ, Inst Logist Sci & Engn, Shanghai 201306, Peoples R China  
Institution Address:  
Shanghai Maritime Univ, Fac Coll Informat Engn, Shanghai 201306, Peoples R China  
Institution:  
Institution:

### 301. IF-USOD: Multimodal information fusion interactive feature enhancement

摘要: Underwater salient object detection (USOD) has garnered increasing attention due to its superior performance in various underwater visual tasks. Despite the growing interest, research on USOD remains in its nascent stages, with existing methods often struggling to capture long-range contextual features of salient objects. Additionally, these methods frequently overlook the complementary nature of multimodal information. The multimodal information fusion can render previously indiscernible objects more detectable, as capturing complementary features from diverse source images enables amore accurate depiction of objects. In this work, we explore an innovative approach that integrates RGB and depth information, coupled with interactive feature enhancement, to advance the detection of underwater salient objects. Our method first leverages the strengths of both transformer and convolutional neural network architectures to extract features from source images. Here, we employ a two-stage training strategy designed to optimize feature fusion. Subsequently, we utilize self-attention and cross-attention mechanisms to model the correlations among the extracted features, thereby amplifying the relevant features. Finally, to fully exploit features across different network layers, we introduce a cross-scale learning strategy to facilitate multi-scale feature fusion, which improves the detection accuracy of underwater salient objects by generating both coarse and fine salient predictions. Extensive experimental evaluations demonstrate the state-of-the-art model performance of our proposed method.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1566-2535  
eISSN:  
1872-6305  
入藏号:  
WOS:001367909900001  
Indexed Date:  
2024-12-07  
Institution Address:  
Shandong Technol & Business Univ, Sch Comp Sci & Technol, Yantai 264000, Peoples R China  
Institution Address:  
Shandong Technol & Business Univ, Shandong Coll & Univ, Coinnovat Ctr, Future Intelligent Comp, Yantai 264000, Peoples R China  
Institution:  
Institution:

### 302. A multi-scale feature fusion convolutional neural network for facial

摘要: This paper designs a new facial Expression recognition network called a multi-scale feature Fusion Convolutional neural Network (EFCN). This network is proposed to solve two problems in the facial expression recognition task. First, there are many commonalities between faces of different expression categories, and the recognition task cannot be precisely performed when the commonality is greater than the individuality. Secondly, facial detail features have a significant impact on the final results of expression recognition, while the image detail features extracted by traditional convolutional neural networks are not sufficient. In order to address the above issues, the feature enhancement network (FEN) and the detail information enhancement module (DEM) are designed. The FEN fuses deep and shallow features. Accordingly, the feature map contains richer information, making it easy to identify the samples. The DEM extracts and fuses the features passed by the backbone network with multi-scale features to enhance the network's ability to extract features from small regions of the face. We validated the proposed method on three datasets, RAF-DB, CK+, and JAFFE, and achieved 84.50%, 97.86%, and 91.05% accuracy, respectively, and the experimental results showed the effectiveness of the proposed method in this paper. For example, on the JAFFE dataset, the recognition accuracy of this method surpasses the MLT method by 1.87%.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0266-4720  
eISSN:  
1468-0394  
入藏号:  
WOS:001126442800001  
Indexed Date:  
2024-01-01  
Institution Address:  
Dalian Nationality Univ, Sch Mech & Elect Engn, Dalian, Peoples R China  
Institution Address:  
Dalian Nationality Univ, Sch Mech & Elect Engn, 18 Liaohe West Rd, Dalian, Peoples R China  
Institution:  
Institution:

### 303. Deep learning-based tool wear prediction and its application for

摘要: Tool wear is a key factor in the cutting process, which directly affects the machining precision and part quality. Accurate tool wear prediction can make proper tool change at an early stage to reduce downtime and enhance product quality. However, traditional methods can not meet the high requirements of the intelligent manufacturing. Therefore, a novel method based on deep learning is proposed to improve the prediction accuracy of tool wear. The multi-scale feature fusion was implemented by the developed parallel convolutional neural networks. The channel attention mechanism combined with the residual connection was developed to consider the weight of the different feature map to enhance the performance of the model. The different tool wear prediction experiments were implemented to verify the superiority of the developed method, and the prediction results of tool wear are more robust and accurate than current methods. Finally, a tool wear monitoring system was developed and applied to the tapping process of the engine cylinder to ensure the quality of the engine cylinder and the stability of the machining process.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
102  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
6  
被引频次合计:  
108  
ISSN:  
0263-2241  
eISSN:  
1873-412X  
入藏号:  
WOS:000649698300005  
Indexed Date:  
2021-06-04  
Institution Address:  
Shanghai Jiao Tong Univ, Sch Mech Engn, Shanghai 200240, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, State Key Lab Mech Syst & Vibrat, Shanghai 200240, Peoples R China  
Institution:  
Institution:

### 304. HmsU-Net: A hybrid multi-scale U-net based on a CNN and transformer for

摘要: Accurate medical image segmentation is of great significance for subsequent diagnosis and analysis. The acquisition of multi-scale information plays an important role in segmenting regions of interest of different sizes. With the emergence of Transformers, numerous networks adopted hybrid structures incorporating Transformers and CNNs to learn multi-scale information. However, the majority of research has focused on the design and composition of CNN and Transformer structures, neglecting the inconsistencies in feature learning between Transformer and CNN. This oversight has resulted in the hybrid network's performance not being fully realized. In this work, we proposed a novel hybrid multi-scale segmentation network named HmsU-Net, which effectively fused multi-scale features. Specifically, HmsU-Net employed a parallel design incorporating both CNN and Transformer architectures. To address the inconsistency in feature learning between CNN and Transformer within the same stage, we proposed the multi-scale feature fusion module. For feature fusion across different stages, we introduced the cross-attention module. Comprehensive experiments conducted on various datasets demonstrate that our approach surpasses current state-of-the-art methods.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
13  
被引频次合计:  
13  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:001173004500001  
PubMed ID:  
38271837  
Indexed Date:  
2024-03-08  
Institution Address:  
Guizhou Univ, Med Coll, Guiyang 550000, Guizhou, Peoples R China  
Institution Address:  
Guizhou Prov Peoples Hosp, Dept Radiol, Key Lab Intelligent Med Imaging Anal & Accurate D, Int Exemplary Cooperat Base Precis Imaging Diag &, Guiyang 550002, Peoples R China  
Institution Address:  
Guizhou Prov Peoples Hosp, Dept Med Imaging, Int Exemplary Cooperat Base Precis Imaging Diag &, 83 Zhongshan East Rd, Guiyang 550002, Guizhou, Peoples R China  
Institution:  
Institution:  
Guizhou Prov Peoples Hosp  
Institution:  
Guizhou Prov Peoples Hosp

### 305. Polycrystalline silicon photovoltaic cell defects detection based on

摘要: In photovoltaic (PV) cell inspection, electroluminescence (EL) imaging provides high spatial resolution for detecting various types of defects. The recent integration of EL imaging with deep learning models has enhanced the recognition of defects in PV cells. However, the high surface impurity content in polycrystalline silicon PV cells presents a challenge. Defect features can be similar to complex background textures, making highlighting features for small target defects difficult and leading to potential misclassification as background or other classes. To address these challenges, we propose a novel deep convolutional neural network (CNN) model for effectively identifying small target defects in polycrystalline PV cells. We first utilize a global context information (GCI) block to improve CNN's modeling of global information, aiding in distinguishing PV cell defects with similar local details. Moreover, we design a channel weight feature pyramid (CWFP) that dynamically adjusts the channel weights of extracted features. We further achieve multi-scale feature fusion through the pyramid structure to enhance the resolution capability for small targets. The proposed model was evaluated on a publicly available dataset of 8 defect classes in polycrystalline PV cells, achieving an accuracy of 96.36 %. The experimental results show that the proposed model outperforms existing deep learning models in detecting small target defects with higher precision.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2352-4928  
入藏号:  
WOS:001331898600001  
Indexed Date:  
2024-10-19  
Institution Address:  
Guilin Univ Elect Technol, Sch Elect Engn & Automat, Guangxi Key Lab Automat Detecting Technol & Instru, Guilin 541004, Peoples R China  
Institution Address:  
Guilin Univ Elect Technol, Sch Architecture & Transportat Engn, Guilin 541004, Peoples R China  
Institution:  
Institution:

### 306. Multi-Scale Feature Fusion Model for Bridge Appearance Defect Detection

摘要: Although the Faster Region-based Convolutional Neural Network (Faster R-CNN) model has obvious advantages in defect recognition, it still cannot overcome challenging problems, such as time-consuming, small targets, irregular shapes, and strong noise interference in bridge defect detection. To deal with these issues, this paper proposes a novel Multi-scale Feature Fusion (MFF) model for bridge appearance disease detection. First, the Faster R-CNN model adopts Region Of Interest (ROI) pooling, which omits the edge information of the target area, resulting in some missed detections and inaccuracies in both detecting and localizing bridge defects. Therefore, this paper proposes an MFF based on regional feature Aggregation (MFF-A), which reduces the missed detection rate of bridge defect detection and improves the positioning accuracy of the target area. Second, the Faster R-CNN model is insensitive to small targets, irregular shapes, and strong noises in bridge defect detection, which results in a long training time and low recognition accuracy. Accordingly, a novel Lightweight MFF (namely MFF-L) model for bridge appearance defect detection using a lightweight network EfficientNetV2 and a feature pyramid network is proposed, which fuses multi-scale features to shorten the training speed and improve recognition accuracy. Finally, the effectiveness of the proposed method is evaluated on the bridge disease dataset and public computational fluid dynamic dataset.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
被引频次合计:  
10  
eISSN:  
2096-0654  
入藏号:  
WOS:001243719700001  
Indexed Date:  
2024-06-25  
Institution Address:  
Southwest Jiaotong Univ, Sch Comp & Artificial Intelligence, Chengdu 611756, Peoples R China  
Institution Address:  
China Merchants Chongqing Rd Engn Inspect Ctr Co, Chongqing 400067, Peoples R China  
Institution Address:  
State Key Lab Bridge Engn Struct Dynam, Chongqing 400067, Peoples R China  
Institution:  
Institution:  
China Merchants Chongqing Rd Engn Inspect Ctr Co  
Institution:  
State Key Lab Bridge Engn Struct Dynam

### 307. MCANet: A joint semantic segmentation framework of optical and SAR

摘要: Deep convolution neural network (DCNN) is among the most effective ways of performing land use classification of high-resolution remote sensing images. Land use classification by fusing optical and synthetic aperture radar (SAR) images has broad application prospects, but related research studies are few. In this study, we developed the first and largest joint optical and SAR land use classification dataset, WHU-OPT-SAR, covering an area of approximately 50,000 km(2), and designed a multimodal-cross attention network (MCANet). MCANet comprises three core modules: the pseudo-siamese feature extraction module, multimodal-cross attention module, and low-high level feature fusion module, which are used for independent feature extraction of optical and SAR images, second-order hidden feature mining, and multi-scale feature fusion. The land use classification accuracy of our approach on the WHU-OPT-SAR dataset was approximately 5% higher than that of optic-image-based approaches. Moreover, the accuracy of city, village, road, water, forest, and farmland classification was improved by 7%, 2%, 5%, 6%, 1%, and 0.6%, respectively, reflecting the superior performance of fusing optical and SAR images. Furthermore, the classification accuracy in Hubei Province of China, which covers an area of 190,000 km(2), has also increased by approximately 5%, which verifies the effectiveness of our approach.  
在中国科学引文数据库中的被引频次:  
8  
在 BIOSIS Citation Index 中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
106  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
116  
ISSN:  
1569-8432  
eISSN:  
1872-826X  
入藏号:  
WOS:000741240400002  
Indexed Date:  
2022-01-18  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & R, Wuhan 430079, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Printing & Packaging, Wuhan 430079, Peoples R China  
Institution:  
Institution:

### 308. BG-3DM2F: Bidirectional gated 3D multi-scale feature fusion for

摘要: A computer-aided diagnosis system is one of the crucial decision support tools under the medical imaging scope. It has recently emerged as a powerful way to diagnose Alzheimer's Disease (AD) from structural magnetic resonance imaging scans. However, due to the deficit of recognition memory in the Mild Cognitive Impairment (MCI) stage, semantic feature ambiguity, and high inter-class visual similarities problems, computer-aided diagnosis of AD remains challenging. To bridge these gaps, this paper proposed a hippocampus analysis method based on a novel 3D convolutional neural network fusion strategy, called Bidirectional Gated 3D Multi-scale Feature Fusion (BG-3DM2F). The suggested BG-3DM2F framework consists of two modules: 3D Multi-Scale Chained Network (3DMS-ChaineNet) and Bidirectional Gated Recurrent Fusion Unit (Bi-GRFU). The 3DMS-ChaineNet architecture is introduced to design the subtle features and capture the variations in hippocampal atrophy, while the Bi-GRFU scheme is investigated to store 3DMS-ChaineNet levels in the forward and backward fashion and retain them in the decision-making process. For validation, our solution is completely evaluated on the public Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. Practically, we conducted empirical evaluations to verify the effect of BG-3DM2F components. In comparison with the current state-of-the-art methods, the experiments show that our proposed approach provides efficient results, achieving the accuracies of 98.12%, 95.26%, and 96.97% for binary classification of Normal Control (NC) versus AD, AD versus MCI, and NC versus MCI, respectively. Therefore, we can conclude that our proposed BG-3DM2F system has the potential to dramatically improve the conventional classification methods for assisting clinical decision-making.  
Web of Science 核心合集中的 "被引频次":  
19  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
19  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:000756332700009  
Indexed Date:  
2022-03-06  
Institution Address:  
Ibn Zohr Univ, Fac Sci, Dept Comp Sci, Lab Comp Syst & Vis LabSIV, BP 8106, Agadir 80000, Morocco  
Institution Address:  
Univ Bordeaux, Bordeaux Lab Comp Sci Res LaBRI UMR 5800, CNRS, Bordeaux INP, F-33400 Talence, France  
Institution Address:  
Univ Victor Segalen Bordeaux 2, Aquitaine Inst Cognit & Integrat Neurosci INCIA U, CNRS, F-33076 Bordeaux, France  
Institution:  
Institution:  
Institution:

### 309. Adaptive Feature Pyramid Networks for Object Detection

摘要: In general object detection, scale variation is always a big challenge. At present, feature pyramid networks are employed in numerous methods to alleviate the problems caused by large scale range of objects in object detection, which makes use of multi-level features extracted from the backbone for top-down upsampling and fusion to acquire a set of multi-scale depth image features. However, the feature pyramid network proposed by Ghiasi et al. adopts a simple fusion method, which fails to consider the fusion feature context, and therefore, it is difficult to acquire good features. In addition, the fusion of multi-scale features directly by traditional upsampling is prone to feature misalignment and loss of details. In this paper, an adaptive feature pyramid network is proposed based on the feature pyramid network to alleviate the foregoing potential problems, which includes two major designs, i.e., adaptive feature upsampling and adaptive feature fusion. The adaptive feature upsampling aims to predict a group of sampling points of each pixel through some models, and constitute feature representation of the pixel by feature combination of sampling points, while adaptive feature fusion is to construct pixel-level fusion weights between fusion features through attention mechanism. The experimental results verified the effectiveness of the method proposed in this paper. On the public object detection dataset MS-COCO test-dev, Faster R-CNN model achieved performance improvement of 1.2 AP by virtue of the adaptive feature pyramid network, and FCOS model could achieve performance improvement of 1.0 AP. What's more, the experiments also validated that the adaptive feature pyramid network proposed herein was more accurate for object localization.  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
42  
被引频次合计:  
44  
ISSN:  
2169-3536  
入藏号:  
WOS:000681239300001  
Indexed Date:  
2021-08-15  
Institution Address:  
Ningbo Univ, Coll Sci & Technol, Ningbo 315300, Peoples R China  
Institution:

### 310. Multi-scale SAR Image Detection Algorithm for Ships Based on Improved

摘要: An multi-scale synthetic aperture radar (SAR) image detection algorithm for ships based on improved YOLOv5 is proposed to address the large pixel scale difference of ship targets in complex scenes and missed detection caused by dense array of ships. For the neck network of YOLOv5, a bi-directional feature pyramid network (BiFPN) is adopted to enhance the multi-scale feature fusion ability of the network, and an enhanced channel-MLP (EC-MLP) module is constructed based on depthwise separable convolution(DSC) and channel MLP in its bottom-up feature fusion branch to enrich semantic information and provide more sufficient ship target context features. The global attention mechanism (GAM) is introduced to enable the network to extract input features selectively and reduce information reduction. In addition, the SIoU loss function is used to further improve the training convergence speed and detection accuracy of the network. Comparative experiments with eight other methods (Faster R-CNN, Libra R-CNN, FCOS, YOLOv5s, PP-YOLOv2, YOLOX-s, PP-YOLOE-s and YOLOv7-tiny) are conducted on SSDD and HRSID datasets. The experimental results show that the AP50 of the improved algorithm reaches 96.7% on SSDD and 95.6% on HRSID, which is superior to the comparison methods.  
摘要:  
针对复杂场景下合成孔径雷达(Synthetic aperture radar, SAR)图像船舶目标像素尺度差异大和船舶密集排列造成目标漏检的问题,提出一种基于改进YOLOv5的船舶多尺度SAR图像检测算法。对于YOLOv5的颈部网络,采用双向特征金字塔结构(Bi-directional feature pyramid network, BiFPN)提升网络多尺度特征融合能力,并在其自下而上的特征融合支路中,基于深度可分离卷积(Depthwise separable convolution, DSC)和通道MLP构建EC-MLP(Enhanced channel-MLP)模块,从而丰富语义信息,提供更充分的船舶目标上下文特征;引入全局注意力机制(Global attention mechanism, GAM),使网络对输入特征进行针对性提取并运算,减少网络的信息丢失;此外,使用SIoU损失函数进一步提高网络的训练收敛速度和检测精度。在SSDD和HRSID数据集上与其他8种方法(Faster R-CNN、Libra R-CNN 、FCOS、YOLOv5s、PP-YOLOv2、YOLOX-s、PP-YOLOE-s和YOLOv7-tiny)进行对比实验。实验结果表明:改进后算法在SSDD数据集上的AP50达到了96.7%,在HRSID数据集上AP50达到了95.6%,优于对比方法。  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1004-9037  
入藏号:  
CSCD:7661157  
Indexed Date:  
2024-05-10  
Institution Address:  
南京邮电大学, 宽带无线通信技术教育部工程研究中心, 南京, 江苏 210003, 中国  
Institution Address:  
Nanjing University of Posts and Telecommunications, Engineering Research Center of the Ministry of Education for Broadband Wireless Communication Technology, 210003  
Institution:  
南京邮电大学  
Institution:  
Nanjing University of Posts and Telecommunications

### 311. UC-former: A multi-scale image deraining network using enhanced

摘要: While convolutional neural networks (CNN) have achieved remarkable performance in single image deraining tasks, it is still a very challenging task due to CNN's limited receptive field and the unreality of the output image. In this paper, UC-former, an effective and efficient U-shaped architecture based on transformer for image deraining was presented. In UC-former, there are two core designs to avoid heavy self-attention computation and inefficient communications across encoder and decoder. First, we propose a novel channel across Transformer block, which computes self-attention between channels. It significantly reduces the computational complexity of high-resolution rain maps while capturing global context. Second, we propose a multi-scale feature fusion module between the encoder and decoder to combine low-level local features and high-level non-local features. In addition, we employ depth-wise convolution and H-Swish non-linear activation function in Transformer Blocks to enhance rain removal authenticity. Extensive experiments indicate that our method outperforms the state-of-the-art deraining approaches on synthetic and real-world rainy datasets.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1077-3142  
eISSN:  
1090-235X  
入藏号:  
WOS:001300054600001  
Indexed Date:  
2024-09-02  
Institution Address:  
Shanghai Maritime Univ, Shanghai 201306, Peoples R China  
Institution:

### 312. M-DETR: Multi-scale DETR for Optical Music Recognition

摘要: Optical Music Recognition (OMR) is an important way to digitize score images and has broad application prospects in fields such as the storage of music documents, music education and digital creation. As a new paradigm for object detection, DETR (detection transformer) has the ability to associate contextual information, which can be exploited to resolve the OMR task. However, the original DETR does not fit OMR well due to its high computational complexity and numerous parameters. To address the DETR defects and improve the recognition accuracy of OMR, we propose a novel multi-scale DETR (M-DETR) with a multi-scale feature fusion mechanism and improved attention mechanisms. First, a new multi-scale feature fusion mechanism is designed to let the backbone network of M-DETR get rich multi-scale information. Then, a key-region attention mechanism is incorporated based on the character that the key information is concentrated on a score image. Finally, the pre-context attention mechanism is introduced to make better use of the contextual association between recognition notes in music scores. Experiment results show that M-DETR achieves recognition accuracy of 90.6% for 7 typical small-sized notes, which is better than Faster R-CNN and YOLO v5, and the improvement rate is 10.02% compared to the original DETR algorithm. The results indicate that M-DETR is an effective way for the OMR task, which also provides a new solution for the detection of small-sized objects with contextual association.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001221608500001  
Indexed Date:  
2024-05-19  
Institution Address:  
East China Univ Sci & Technol, Shanghai, Peoples R China  
Institution Address:  
Univ Bio Bio, Dept Comp Sci & Informat Technol, Chillan 3780000, Chile  
Institution Address:  
Shanghai Key Lab Comp Software Evaluating Testing, Shanghai, Peoples R China  
Institution:  
Institution:  
Institution:  
Shanghai Key Lab Comp Software Evaluating Testing

### 313. A Few-Shot Learning Method for SAR Images Based on Weighted Distance and

摘要: Convolutional Neural Network (CNN) has been widely applied in the field of synthetic aperture radar (SAR) image recognition. Nevertheless, CNN-based recognition methods usually encounter the problem of poor feature representation ability due to insufficient labeled SAR images. In addition, the large inner-class variety and high cross-class similarity of SAR images pose a challenge for classification. To alleviate the problems mentioned above, we propose a novel few-shot learning (FSL) method for SAR image recognition, which is composed of the multi-feature fusion network (MFFN) and the weighted distance classifier (WDC). The MFFN is utilized to extract input images' features, and the WDC outputs the classification results based on these features. The MFFN is constructed by adding a multi-scale feature fusion module (MsFFM) and a hand-crafted feature insertion module (HcFIM) to a standard CNN. The feature extraction and representation capability can be enhanced by inserting the traditional hand-crafted features as auxiliary features. With the aid of information from different scales of features, targets of the same class can be more easily aggregated. The weight generation module in WDC is designed to generate category-specific weights for query images. The WDC distributes these weights along the corresponding Euclidean distance to tackle the high cross-class similarity problem. In addition, weight generation loss is proposed to improve recognition performance by guiding the weight generation module. Experimental results on the Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset and the Vehicle and Aircraft (VA) dataset demonstrate that our proposed method surpasses several typical FSL methods.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
16  
被引频次合计:  
16  
eISSN:  
2072-4292  
入藏号:  
WOS:000859920800001  
Indexed Date:  
2022-10-05  
Institution Address:  
Beihang Univ, Sch Elect & Informat Engn, Beijing 100191, Peoples R China  
Institution Address:  
Edinburgh Napier Univ, Cyber & Big Data Res Lab, Edinburgh EH11 4BN, Midlothian, Scotland  
Institution Address:  
Univ Leicester, Dept Informat, Leicester LE1 7RH, Leics, England  
Institution:  
Institution:  
Institution:

### 314. Predicting Malignancy and Benign Thyroid Nodule Using Multi-Scale

摘要: Nowadays, thyroid ultrasound examination faces some problems such as weak effective feature information, plentiful noise, and small samples. Our research aims at helping doctors making decision more accurately and quickly to identify the characteristics of patients' thyroid nodules based on ultrasound images. Firstly, after pre-processing ultrasound images of thyroid nodules, a noise reduction method is proposed by using weighted adaptive gamma correction which can effectively suppress the generation of noise and improve the global information contrast ratio. Secondly, fine-tuning transfer learning to pre-train ResNet-18 convolutional neural network is used to solve over-fitting under small samples. Thirdly, an adaptive threshold Local Ternary Pattern algorithm is proposed to extract local texture features of the ultrasound images in order to enhance the classification performance. Finally, a multi-scale feature fusion approach, which combines the local texture features and the deep features (the global texture features) automatically extracted by convolutional layers, is carried out by following a second fine-tuning training in ResNet-18 convolutional neural network based on the multi-scale joint features. The test results show: (1) the improved Adaptive Threshold Local Ternary Pattern algorithm demonstrates better performance than other algorithms in extracting texture features on the experimental thyroid nodule dataset, which has fewer misclassified samples and can better describe the texture information of the ultrasound image nodules. (2) The classification accuracy is significantly promoted in the given real test set based on the improving ResNet-18 convolutional neural network by using the proposed multi-scale feature fusion approach.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1054-6618  
eISSN:  
1555-6212  
入藏号:  
WOS:000734794700023  
Indexed Date:  
2022-01-03  
Institution Address:  
Nanchang Univ, Dept Sch Qianhu, Xuefu Rd, Nanchang 330031, Jiangxi, Peoples R China  
Institution Address:  
Nanchang Univ, Dept Comp, Xuefu Rd, Nanchang 330031, Jiangxi, Peoples R China  
Institution Address:  
Nanchang Univ, Dept Ultrasound, Affiliated Hosp 1, Nanchang 330031, Jiangxi, Peoples R China  
Institution:  
Institution:  
Institution:

### 315. Single-Core Multiscale Residual Network for the Super Resolution of

摘要: In a gravity-free or microgravity environment, liquid metals without crystalline nuclei achieve a deep undercooling state. The resulting melts exhibit unique properties, and the research of this phenomenon is critical for exploring new metastable materials. Owing to the rapid crystallization rates of deeply undercooled liquid metal droplets, as well as cost concerns, experimental systems meant for the study of liquid metal specimens usually use low-resolution, high-framerate, high-speed cameras, which result in low-resolution photographs. To facilitate subsequent studies by material scientists, it is necessary to use super-resolution techniques to increase the resolution of these photographs. However, existing super-resolution algorithms cannot quickly and accurately restore the details contained in images of deeply undercooled liquid metal specimens. To address this problem, we propose the single-core multiscale residual network (SCMSRN) algorithm for photographic images of liquid metal specimens. In this model, multiple cascaded filters are used to obtain feature information, and the multiscale features are then fused by a residual network. Compared to existing state-of-the-art artificial neural network super-resolution algorithms, such as SRCNN, VDSR and MSRN, our model was able to achieve higher PSNR and SSIM scores and reduce network size and training time.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2504-4990  
入藏号:  
WOS:000913080700001  
Indexed Date:  
2021-06-01  
Institution Address:  
Beijing Jiaotong Univ, Inst Photoelect Technol, Sch Sci, Beijing 102603, Peoples R China  
Institution Address:  
North China Univ Technol, Sch Comp Sci & Technol, Beijing 100144, Peoples R China  
Institution:  
Institution:

### 316. Infrared Small Target Detection Method with Vision Transformer and Dual

摘要: The existing infrared small-target detection method based on convolutional neural networks (CNN) exhibits the problem of a limited receptive field in the encoder stage, and the decoder lacks an effective feature interaction when fusing multiscale features. To address the aforementioned issues, in this study, a new method is proposed based on an encoder-decoder structure. Specifically, a vision transformer is used as an encoder to extract multiscale features from small infrared target images. The vision transformer is an emerging deeplearning architecture that uses a self-attention mechanism to capture the global relationship between all pixels in the input image, thereby effectively processing long-range dependencies and contextual information in the image. Furthermore, a dual-decoder module, comprising an interactive decoder and auxiliary decoder, is proposed to improve the ability of the decoder to reconstruct small infrared targets. The dual-decoder module can make full use of the complementary information between different features, promote interaction between deep and shallow features, and better reconstruct small infrared targets by combining the results of the two decoders. Experimental results on widely used public datasets show that the proposed method outperforms other methods in terms of two evaluation indicators: F\_1 and mIoU.  
摘要:  
当前基于卷积神经网络的红外小目标检测方法在编码器阶段受限于感受野,且解码器在多尺度特征融合中缺乏有效的特征交互。本文提出了一种基于编码器-解码器结构的新方法,针对现有红外小目标检测方法中的问题进行改进。该方法使用视觉Transformer作为编码器,能够有效地提取红外小目标图像的多尺度特征。视觉Transformer是一种新兴的深度学习架构,其通过自注意力机制捕捉图像中像素之间的全局关系,以处理长程依赖性和上下文信息。此外,本文还设计了一个由交互式解码器和辅助解码器组成的双解码器模块,旨在提高解码器对红外小目标的重构能力。该双解码器模块能够充分利用不同特征之间的互补信息,促进深层特征和浅层特征之间的交互,并通过将两个解码器的结果进行叠加,以更好地重构红外小目标。在广泛使用的公共数据集上的实验结果表明,本文提出的方法在F\_1和mIoU两个评价指标上的性能优于其他对比方法。  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1001-8891  
入藏号:  
CSCD:7806658  
Indexed Date:  
2024-11-23  
Institution Address:  
重庆邮电大学通信与信息工程学院, 重庆 400065, 中国  
Institution Address:  
重庆理工大学电气与电子工程学院, 重庆 400054, 中国  
Institution Address:  
Department of Communication and Information Engineering,Chongqing University of Posts and Telecommunications, Chongqing 400065  
Institution Address:  
Department of Electrical and Electronic Engineering,Chongqing University of Technology, Chongqing 400054  
Institution:  
重庆邮电大学通信与信息工程学院  
Institution:  
重庆理工大学电气与电子工程学院  
Institution:  
Department of Communication and Information Engineering,Chongqing University of Posts and Telecommunications  
Institution:  
Department of Electrical and Electronic Engineering,Chongqing University of Technology

### 317. Real-time Detection of Surface Cracks in Steel Beam using Lightweight

摘要: Steel beams are widely used in engineering projects,and if their surface cracks are not found in time,they may bring potential safety hazards.In this paper,a lightweight convolutional neural network with a cross-stage hierarchical structure is used to achieve rapid real-time detection of surface cracks in steel beam.Firstly,a cross-stage local network to build a backbone network for feature extraction is used,which not only enriches the gradient update path,but also helps extract the shallow features of cracks.Secondly,the cross-stage layering module is embedded as a feature extractor into one of the branches of the cross-stage layered structure to obtain a lightweight feature extraction module,which greatly improves the speed of crack detection.Finally,the multiscale feature fusion will be combined with the YOLO layer to complete the object detection.Experiments show that the highest mAP of the lightweight convolutional neural network with a cross-stage layered structure is 93.59%,and the frame rate is 30.3 s~(-1).Under the premise that the detection performance gap is not big,its detection speed is 4 times faster than YOLOv3 and 4.5 times faster than YOLOv4.  
摘要:  
针对钢梁在工程项目中应用广泛,其表面缺陷若未能及时发现将很可能带来安全隐患。本文利用一种具有跨阶段分层结构的轻量化卷积神经网络实现了钢梁表面缺陷的快速实时检测。首先使用跨阶段局部网络搭建用于特征提取的骨干网络,不仅能丰富了梯度更新路径,而且有助于浅层表面缺陷特征的提取。其次,将跨阶段分层模块作为特征提取器嵌入到跨阶段分层结构的其中一个分支中得到轻量化的特征提取模块,极大的提高了检测速度。最后,将多尺度特征融合与YOLO层相结合完成目标检测任务。实验表明,具有跨阶段分层结构的轻量化卷积神经网络最高mAP为93.59%,帧率为30.3 s~(-1)。在检测性能差距不大的前提下,其检测速度较YOLOv3提高了4倍,与YOLOv4相比提高了4.5倍。  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1003-8728  
入藏号:  
CSCD:7209223  
Indexed Date:  
2022-07-22  
Institution Address:  
河南省工业学校机电工程系, 郑州, 河南 450011, 中国  
Institution Address:  
昆明理工大学机电工程学院, 昆明, 云南 650500, 中国  
Institution Address:  
Department of Mechanical and Electrical Engineering,Industrial School of Henan Province, Zhengzhou, Henan 450011, China  
Institution Address:  
Faculty of Mechanical and Electrical Engineering,Kunming University of Science and Technology, Kunming, Yunnan 650500, China  
Institution:  
河南省工业学校机电工程系  
Institution:  
昆明理工大学机电工程学院  
Institution:  
Department of Mechanical and Electrical Engineering,Industrial School of Henan Province  
Institution:  
Faculty of Mechanical and Electrical Engineering,Kunming University of Science and Technology

### 318. An Explainable Deep Learning Framework for Sorghum Weed Classification

摘要: Sorghum is a vital crop known for its resilience in harsh environmental conditions, playing a key role in food security and bioenergy production across the globe. Accurate sorghum weed classification is crucial for efficient agricultural management, yet traditional methods are labour-intensive and resource-heavy. This research presents the first work on the sorghum weed classification dataset, introducing a novel deep-learning model using a customised DenseNet-169, enhanced with two innovative feature extraction blocks: Dimensional Convolutional Feature Fusion (DCFF) and Deformable Multi-Kernel Channel Fusion (DMKCF). These blocks improve multi-scale feature extraction and adaptively capture complex patterns from the weed images. The model also integrates Triplet Attention (TA) to refine feature representations and employs Grad-CAM and LIME for visual interpretability, offering insights into the classification process. The proposed network, when tested on the SorghumWeedDataset\_Classification, achieved a testing accuracy of 99.07%, surpassing several state-of-the-art CNN models. The proposed model provides a robust solution for sorghum weed classification, advancing agricultural monitoring and management practices. Additionally, the model achieved 100% accuracy for the broad leaf weed class, further underscoring its effectiveness in identifying and classifying sorghum weeds with absolute precision.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001422034400009  
Indexed Date:  
2025-02-21  
Institution Address:  
Vellore Inst Technol, Sch Comp Sci & Engn SCOPE, Chennai 600127, India  
Institution:

### 319. Improved multi-scale flame detection method

摘要: The deepening of the number of network layers can weaken the ability to characterize the detailed information of the deep features of the flame target, and at the same time extract redundant features with low correlation, resulting in low flame recognition accuracy. Aiming at this problem, a flame detection method based on improved Faster R-CNN is proposed to improve the accuracy of flame recognition in deep networks. Firstly, the ResNet50 network is used to extract flame features, and the SENet module is added to reduce the redundant features of flame targets. Then, the deep features and shallow features are multi-scale feature fusion to enhance the detailed information of deep features. Finally, the network is trained to realize the recognition of flame targets positioning. In the experiment, the VOC flame data set is constructed for network training, the test set is used for detection, and the feature map visualization is compared. Compared with the model before the improvement, the AP value increases by 7.78 , the recall increases by 9.05% , and the precision increases by 12.54%. By combining the attention mechanism module and the multi-scale feature fusion mechanism, the flame target detection model proposed in this paper, can effectively extract the flame target feature, and the flame target detection result is more accurate.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1007-2780  
eISSN:  
2097-3217  
入藏号:  
WOS:000656759500015  
Indexed Date:  
2021-06-24  
Institution Address:  
Xian Univ Architecture & Technol, Coll Informat & Control Engn, Xian 710055, Peoples R China  
Institution:

### 320. Research on Marine Ship Detection Based on Multi-scale Feature Fusion

摘要: In order to enhance the safety of maritime traffic, an improved ship detection algorithm MS-YOLOV3 is proposed to solve the problems of detection based on YOLOV3 on the ship data set, such as the false detection of target frame and the omission of small targets and low detection accuracy. Firstly, the ship image data set is constructed, including data collection, enhancement and label labeling. The dimension clustering algorithm is used to find the anchor box of appropriate size in the data set and applied to the corresponding scale feature map. Secondly, the model extracts network features based on Darknet-53 and increases network prediction scale. The DCA feature fusion strategy is added to maximize the correlation between high-level features and low-level features, and then the channel splicing is carried out. Measures as said above really improve the detection ability of the model. Finally, GIOU is used as both the parameter of the border optimization function and the border screening index to improve the accuracy of the network in predicting the position information of the border box. The experimental results of MS-YOLOV3 and YOLOV3 show that the accuracy of the former is improved by 7.9 percentage points. At the same time, the added loss of GIOU frame reduces the average loss of the model, strengthens the robustness of the model, and greatly reduces the positioning error of the target frame. According to the training effect on Pascal VOC2007 data set, the average accuracy of MS-YOLOV3 is better than the series algorithm of YOLO, SSD300 and Faster-RCNN. The proposed feature interactive detection model in this paper makes the ship position information and classification more accurate.  
摘要:  
为了加强海上交通的安全性,以常见的民用船和军用船为研究对象,针对原始YOLOV3算法在船舶数据集上检测精度不高、目标框出现误检和小目标漏检的问题,提出了改进的船舶检测算法MS-YOLOV3。构建船舶图像数据集Shipdataset,包括数据采集、增强和标签标注,使用维度聚类算法在该数据集中找出合适尺寸的先验框,并应用于相对应的尺度特征图。以Darknet-53的网络框架为基础特征提取网络,增加网络预测尺度,在多尺度特征融合中加入DCA融合策略,提高模型对船舶的检测能力。以MS-YOLOV3为算法框架,采用GIOU作为边框损失函数的参数,提升模型对边界框位置信息的预测准确度。结果MS-YOLOV3与YOLOV3检测算法的对比实验表明,前者在船舶数据集上的精度有7.9个百分点的提升。同时加入的GIOU边框损失,拉低了模型的平均损失,加强了模型的鲁棒性,使得目标框的定位误差大大减小。根据Pascal VOC2007数据集上的训练效果,MS-YOLOV3的平均精度相较于YOLO系列算法、SSD300和Faster-RCNN,精确度优势更加明显。提出的MS-YOLOV3检测模型使得船舶的位置信息和类别精度更加准确。  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1002-8331  
入藏号:  
CSCD:7171664  
Indexed Date:  
2022-05-20  
Institution Address:  
江苏科技大学计算机学院, 镇江, 江苏 212003, 中国  
Institution Address:  
School of Computer Science, Jiangsu University of Science and Technology, Zhenjiang, Jiangsu 212003, China  
Institution:  
江苏科技大学计算机学院  
Institution:  
School of Computer Science, Jiangsu University of Science and Technology

### 321. Cardiac Magnetic Resonance Image Segmentation Method Based on

摘要: Accurate segmentation of the left atrial structure using magnetic resonance images provides an important basis for the diagnosis of atrial fibrillation (AF) and its treatment using robotic surgery. In this study, an image segmentation method based on sequence relationship learning and multi-scale feature fusion is proposed for 3D to 2D sequence conversion in cardiac magnetic resonance images and the varying scales of left atrial structures within different slices. Firstly, a convolutional neural network layer with an attention module was designed to extract and fuse contextual information at different scales in the image, to strengthen the target features using the correlation between features in different regions within the image, and to improve the network's ability to distinguish the left atrial structure. Secondly, a recurrent neural network layer oriented to two-dimensional images was designed to capture the correlation of left atrial structures in adjacent slices by simulating the continuous relationship between sequential image slices. Finally, a combined loss function was constructed to reduce the effect of positive and negative sample imbalance and improve model stability. The Dice, IoU, and Hausdorff distance values reached 90.73%, 89.37%, and 4.803 mm, respectively, based on the LASC2013 (left atrial segmentation challenge in 2013) dataset; the corresponding values reached 92.05%, 89.41% and 9.056 mm, respectively, based on the ASC2018 (atrial segmentation challenge at 2018) dataset.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
1424-8220  
入藏号:  
WOS:000916420200001  
PubMed ID:  
36679487  
Indexed Date:  
2023-02-09  
Institution Address:  
Donghua Univ, Coll Mech Engn, Shanghai 201620, Peoples R China  
Institution Address:  
Donghua Univ, Inst Artificial Intelligence, Shanghai 201620, Peoples R China  
Institution:  
Institution:

### 322. Pose estimation algorithm for robot arm target grabbing based on RGB

摘要: Aiming at the problem that the target object was partially occluded during the capture process,a deep convolutional neural network based on multi-scale feature fusion was designed to extract the projection feature points of the three-dimensional target,and the different Perspective-n-Point(PnP)algorithm was used. The network uses synthetic data automatically generated by the computer for training. After verification, the network trained with synthetic data could also work effectively in real scenes. Finally,a UR5 robotic arm grasping platform was built based on Robot Operating System(ROS). The trained model was deployed on the platform for grasping experiments to verify the actual application effect of the pose estimation algorithm. Experiments showed that the proposed method using only RGB image information could estimate the pose of the target object, and capture the object with unknown pose in the actual scene.  
摘要:  
针对抓取过程中目标物体部分被遮挡的问题,设计了一种基于多尺度特征融合的深度卷积神经网络提取3D目标的投影特征点,并根据不同投影特征点数采用不同的多点透视成像算法。网络使用计算机自动生成的合成数据进行训练,经过验证,使用合成数据训练的网络也能在真实场景中有效工作。最后,搭建了一个基于机器人操作系统的UR5机械臂抓取平台,将训练好的模型部署到该平台上进行抓取实验,结果表明所提方法能够估计出目标物体的位姿,并在实际场景中抓取位姿未知的物体。  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-5911  
入藏号:  
CSCD:7585198  
Indexed Date:  
2024-01-12  
Institution Address:  
广西大学电气工程学院;;广西大学电气工程学院, 广西制造系统与先进制造技术重点实验室;;广西电力装备智能控制与运维重点实验室, 南宁;;南宁, 广西;;广西 530004;;530004, 中国  
Institution Address:  
广西大学电气工程学院, 广西电力装备智能控制与运维重点实验室, 南宁, 广西 530004, 中国  
Institution Address:  
School of Electrical Engineering,Guangxi University;;School of Electrical Engineering,Guangxi University, Guangxi Key Laboratory of Manufacturing System & Advanced Manufacturing Technology;;Guangxi Key Laboratory of Intelligent Control and Maintenance of Power Equipment, Nanning;;Nanning, ;; 530004;;530004  
Institution Address:  
School of Electrical Engineering,Guangxi University, Guangxi Key Laboratory of Intelligent Control and Maintenance of Power Equipment, Nanning, Guangxi 530004, China  
Institution:  
广西大学电气工程学院;;广西大学电气工程学院  
Institution:  
广西大学电气工程学院  
Institution:  
School of Electrical Engineering,Guangxi University;;School of Electrical Engineering,Guangxi University  
Institution:  
School of Electrical Engineering,Guangxi University

### 323. Multi-scale residual network model combined with Global Average Pooling

摘要: Human Action Recognition is a research hotspot in the field of computer vision. However, due to the complexity of the environment and the diversity of actions, Human Action Recognition still faces many challenges. At the same time, traditional CNN has problems such as single feature scale, decreased accuracy of deep network, and excessive network parameters. Aiming at the above research problems, this paper proposes a novel residual network model based on Multi-scale Feature Fusion and Global Average Pooling. The model uses a Multi-scale Feature Fusion module to extract feature information of different scales, enriches spatial-time information. At the end of the network, Global Average Pooling is used to instead of a Fully Connected layer. Compared with a Fully Connected layer, Global Average Pooling will dilute the combination of the relative positions of different features. Therefore, the features trained by convolution are more effective. In addition, Global Average Pooling can realize direct mapping between output channels and feature categories to reduce excessive model parameters. The model in this paper is verified on the UT-interaction dataset, UCF11 (YouTube Action dataset), UCF101 dataset and CAVIAR dataset. The results show that compared with the state-of-the-art approaches, this approach has high recognition accuracy and excellent robustness, and has excellent performance on datasets with complex backgrounds and diverse action categories.  
在 BIOSIS Citation Index 中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
被引频次合计:  
15  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:000702205000001  
Indexed Date:  
2022-06-15  
Institution Address:  
Inner Mongolia Univ Sci & Technol, Sch Elect & Informat Engn, 7 Arding St, Baotou, Peoples R China  
Institution:

### 324. A Lightweight Object Counting Network Based on Density Map Knowledge

摘要: Object counting aims to count the accurate number of object instances in images, and its operation efficiency is essential. However, most current CNN-based methods rely on complex network architectures, which results in them consuming a significant amount of memory, time, and other resources at runtime. This seriously limits their deployment in practical application scenarios, such as public safety and agriculture planting. Therefore, we propose a lightweight object counting method named EdgeCount to effectively balance inference speed and object counting accuracy. Specifically, we construct a network composed of a student model (EdgeCount) and a teacher model (EdgeCount-T) with the same encoder-decoder structure based on density map knowledge distillation (DMKD), allowing the EdgeCount to learn object density distribution from the EdgeCount-T. After that, we introduce spatial and channel reconstruction convolution (SCConv), composed of a spatial reconstruction unit (SRU) and a channel reconstruction unit (CRU), to decrease spatial and channel redundancy with lower computational costs. Moreover, a low parameter weighted multi-scale feature fusion module (LWMFFM) is designed to further improve the countering ability through segmenting minor structural discrepacies among multi-scale features. Extensive experiments conducted on challenging remote sensing and dense crowd object counting datasets demonstrate the effectiveness and superiority of our method. In particular, under the four NVIDIA Jetson devices, EdgeCount can accurately counter objects with only 0.12M parameters and 19.87M floating-point operations per second (FLOPs) in the size of 128, which achieves the lowest latency and fastest FPS compared with other state-of-the-art object counters.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1051-8215  
eISSN:  
1558-2205  
入藏号:  
WOS:001422045800043  
Indexed Date:  
2025-02-23  
Institution Address:  
Chongqing Univ Posts & Telecommun, Sch Commun & Informat Engn, Chongqing 400065, Peoples R China  
Institution Address:  
Xidian Univ, Sch Elect Engn, State Key Lab Integrated Serv Networks, Xian 710071, Shaanxi, Peoples R China  
Institution Address:  
Brunel Univ London, Dept Elect & Elect Engn, London UB8 3PH, England  
Institution:  
Institution:  
Institution:

### 325. Crowd Counting Based on Multiscale Spatial Guided Perception Aggregation

摘要: Crowd counting has received extensive attention in the field of computer vision, and methods based on deep convolutional neural networks (CNNs) have made great progress in this task. However, challenges such as scale variation, nonuniform distribution, complex background, and occlusion in crowded scenes hinder the performance of these networks in crowd counting. In order to overcome these challenges, this article proposes a multiscale spatial guidance perception aggregation network (MGANet) to achieve efficient and accurate crowd counting. MGANet consists of three parts: multiscale feature extraction network (MFEN), spatial guidance network (SGN), and attention fusion network (AFN). Specifically, to alleviate the scale variation problem in crowded scenes, MFEN is introduced to enhance the scale adaptability and effectively capture multiscale features in scenes with drastic scale variation. To address the challenges of nonuniform distribution and complex background in population, an SGN is proposed. The SGN includes two parts: the spatial context network (SCN) and the guidance perception network (GPN). SCN is used to capture the detailed semantic information between the multiscale feature positions extracted by MFEN, and improve the ability of deep structured information exploration. At the same time, the dependence relationship between the spatial remote context is established to enhance the receptive field. GPN is used to enhance the information exchange between channels and guide the network to select appropriate multiscale features and spatial context semantic features. AFN is used to adaptively measure the importance of the above different features, and obtain accurate and effective feature representations from them. In addition, this article proposes a novel region-adaptive loss function, which optimizes the regions with large recognition errors in the image, and alleviates the inconsistency between the training target and the evaluation metric. In order to evaluate the performance of the proposed method, extensive experiments were carried out on challenging benchmarks including ShanghaiTech Part A and Part B, UCF-CC-50, UCF-QNRF, and JHU-CROWD++. Experimental results show that the proposed method has good performance on all four datasets. Especially on ShanghaiTech Part A and Part B, CUCF-QNRF, and JHU-CROWD++ datasets, compared with the state-of-the-art methods, our proposed method achieves superior recognition performance and better robustness.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
2162-237X  
eISSN:  
2162-2388  
入藏号:  
WOS:001373104900029  
PubMed ID:  
37610898  
Indexed Date:  
2024-12-28  
Institution Address:  
Hangzhou Dianzi Univ, Dept Automat, Hangzhou 310018, Peoples R China  
Institution:

### 326. Design of Neural Network Model for Cross-Media Audio and Video Score

摘要: In this paper, the residual convolutional neural network is used to extract the note features in the music score image to solve the problem of model degradation; then, multiscale feature fusion is used to fuse the feature information of different levels in the same feature map to enhance the feature representation ability of the model. A network composed of a bidirectional simple loop unit and a chained time series classification function is used to identify notes, parallelizing a large number of calculations, thereby speeding up the convergence speed of training, which also makes the data in the dataset no longer need to be strict with labels. Alignment also reduces the requirements on the dataset. Aiming at the problem that the existing cross-modal retrieval methods based on common subspace are insufficient for mining local consistency within modalities, a cross-modal retrieval method fused with graph convolution is proposed. The K-nearest neighbor algorithm is used to construct modal graphs for samples of different modalities, and the original features of samples from different modalities are encoded through a symmetric graph convolutional coding network and a symmetric multilayer fully connected coding network, and the encoded features are fused and input. We jointly optimize the intramodal semantic constraints and intermodal modality-invariant constraints in the common subspace to learn highly locally consistent and semantically consistent common representations for samples from different modalities. The error value of the experimental results is used to illustrate the effect of parameters such as the number of iterations and the number of neurons on the network. In order to more accurately illustrate that the generated music sequence is very similar to the original music sequence, the generated music sequence is also framed, and finally the music sequence spectrogram and spectrogram are generated. The accuracy of the experiment is illustrated by comparing the spectrogram and the spectrogram, and genre classification predictions are also performed on the generated music to show that the network can generate music of different genres.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1687-5265  
eISSN:  
1687-5273  
入藏号:  
WOS:000823736700011  
PubMed ID:  
35733575  
Indexed Date:  
2022-07-21  
Institution Address:  
Xinyang Normal Univ, Xinyang 464000, Henan, Peoples R China  
Institution:

### 327. SLG-Net: Small-Large-Global Feature-Based Multilevel Feature Extraction

摘要: Automatic ultrasound image segmentation improves the efficiency of clinical diagnosis and decreases the workload of doctors. Many ultrasound image segmentation methods only focus on capturing local details and global dependencies, whereas ignoring large-scale context information. However, it is essential to extract large-scale context features for large targets in images. To enhance the capability of feature extraction of the model for targets with various sizes and improve segmentation performance, we propose an effective multilevel feature extraction network (SLG-Net) which can extract features from local small details, large-scale context to global dependencies. The SLG-Net is parallel dual-encoder architecture which consists of a CNN encoder and a transformer encoder. Specifically, the CNN encoder improves the representation and interaction of fine feature and large-scale context feature for targets of different sizes by large-small kernel attention (LSKA) modules. The LSKA module firstly extracts features by parallel small kernel module and large-scale feature selection (LSFS) module. The extracted features from above modules are added for further information interaction through a following multi-scale feature interaction module. To fully leverage the feature extraction capability of large kernel convolutions and decrease the number of parameters, we design the large kernel decomposition module (LKDM) to extract large-scale context features in LSFS module. The transformer encoder is used to capture global features for compensating the limitations of CNN encoder. To merge multilevel features, a multi-scale feature fusion module is introduced after the dual-encoder. In addition, at the skip connection, a multi-scale attention module is integrated to retain significant shallow features for subsequent fusion of deep and shallow features. Experiments on three public ultrasound datasets indicate that the proposed network accomplishes the prominent performance for ultrasound image segmentation. It shows the potential of our study to promote intelligence in clinical medicine.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001403265600028  
Indexed Date:  
2025-01-30  
Institution Address:  
Changsha Univ Sci & Technol, Sch Elect & Informat Engn, Changsha 410114, Peoples R China  
Institution Address:  
Changsha Univ Sci & Technol, State Key Lab Disaster Prevent & Reduct Power Grid, Changsha 410114, Peoples R China  
Institution Address:  
Cent South Univ, Xiangya Hosp, Dept Neurol, Changsha 410008, Peoples R China  
Institution Address:  
Cent South Univ, Xiangya Hosp, Natl Clin Res Ctr Geriatr Disorders, Changsha 410008, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 328. Attention-Guided Feature Extraction and Multiscale Feature Fusion 3D

摘要: Automatic detection of pulmonary nodules is critical for the early diagnosis and prevention of lung cancer. Computed tomography (CT) is an effective and economical lung cancer detection method. In CT images, the size and shape of pulmonary nodules appear different, and some nodules appear similar to the surrounding tissues. Therefore, the automatic localization of pulmonary nodules in CT images is a challenging task. An attention-embedded three-dimensional convolutional neural network is proposed for pulmonary nodule detection in the current study. Specifically, 1) channel-spatial attention guides 3D ResNet to down sample the input 3D CT patch. The channel pays attention to important features and the space to the region of interest. The two form a complementary feature extraction mechanism to effectively help the global flow of information in the network and refine the feature mapping to extract the nodule context features. 2) The channel-spatial attention module changes the fusion model of the feature pyramid, adaptively adjusts the pixel-level weight between features and extracts multi-scale representative node features. 3) The deep separable convolution is used to replace the standard convolution of ResNet, reducing the time cost and improving the efficiency of model training on the premise of ensuring the model's performance. 4) To adapt the distribution of nodule scale, different characteristic layers correspond to two sizes of anchors. Under the condition of ensuring the detection rate of nodules, the number of anchor frames is reduced, and the network sensitivity is improved. Finally, several ablation experiments are carried out using the LUNA16 dataset. The results revealed that the attention-guided network could extract the multi-scale representative features of nodules, and the average sensitivity was 97.7%. Additionally, the CMP score reached 0.912. The extensive experiments demonstrate that the proposed approach can effectively improve the detection sensitivity and control the number of false positive nodules, which has clinical application value and a certain reference value.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
2169-3536  
入藏号:  
WOS:000811545500001  
Indexed Date:  
2022-06-24  
Institution Address:  
Army Engn Univ PLA, Nanjing 210007, Peoples R China  
Institution Address:  
Hainan Normal Univ, Haikou 571158, Hainan, Peoples R China  
Institution:  
Institution:

### 329. Swin-Net: A Swin-Transformer-Based Network Combing with Multi-Scale

摘要: Breast cancer is one of the most common cancers in the world, especially among women. Breast tumor segmentation is a key step in the identification and localization of the breast tumor region, which has important clinical significance. Inspired by the swin-transformer model with powerful global modeling ability, we propose a semantic segmentation framework named Swin-Net for breast ultrasound images, which combines Transformer and Convolutional Neural Networks (CNNs) to effectively improve the accuracy of breast ultrasound segmentation. Firstly, our model utilizes a swin-transformer encoder with stronger learning ability, which can extract features of images more precisely. In addition, two new modules are introduced in our method, including the feature refinement and enhancement module (RLM) and the hierarchical multi-scale feature fusion module (HFM), given that the influence of ultrasonic image acquisition methods and the characteristics of tumor lesions is difficult to capture. Among them, the RLM module is used to further refine and enhance the feature map learned by the transformer encoder. The HFM module is used to process multi-scale high-level semantic features and low-level details, so as to achieve effective cross-layer feature fusion, suppress noise, and improve model segmentation performance. Experimental results show that Swin-Net performs significantly better than the most advanced methods on the two public benchmark datasets. In particular, it achieves an absolute improvement of 1.4-1.8% on Dice. Additionally, we provide a new dataset of breast ultrasound images on which we test the effect of our model, further demonstrating the validity of our method. In summary, the proposed Swin-Net framework makes significant advancements in breast ultrasound image segmentation, providing valuable exploration for research and applications in this domain.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
7  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
eISSN:  
2075-4418  
入藏号:  
WOS:001160149100001  
PubMed ID:  
38337784  
Indexed Date:  
2024-02-19  
Institution Address:  
Cent South Univ, Sch Humanities, Changsha 410012, Peoples R China  
Institution Address:  
Cent South Univ, Sch Comp Sci & Engn, Changsha 410083, Peoples R China  
Institution Address:  
Cent South Univ, Affiliated Canc Hosp, Xiangya Sch Med, Dept Med Ultrasound,Hunan Canc Hosp, Changsha 410031, Peoples R China  
Institution:  
Institution:  
Institution:

### 330. Prediction model of theft crime based on the dynamic fusion of

摘要: A prediction model combining self-attention and dynamic fusion of multiscale and multiview features is proposed to solve the problems of inaccurate fusion of spatiotemporal prediction features and insufficient temporal dynamic adaptability of theft crime. Initially, data are processed by constructing a method that can match case data with different lengths of time series to an adaptive length by projecting the crime data onto the map grid based on local longitude and latitude information. After word vector mapping, the weather, crime time, and location are used to construct the input vector of multidimensional feature fusion. In addition, a self-attention mechanism is introduced to generate the vector of a dynamic fusion of multiview features. The final step involves encoding the dynamic fusion vector of perspective features and sending it to the classifier to predict the crime situation in each map grid. By validating the method on a real dataset of theft crimes in a city, the proposed model can achieve a maximum prediction precision of 0.899 at three different geographic grid divisions, which is significantly better than other comparable models.  
摘要:  
针对盗窃犯罪时空预测特征融合不精、时序动态适应性不足问题,提出自注意力和多尺度多视角特征动态融合的预测模型。首先,以盗窃发案的位置信息为基础,将数据投射到地图栅格内,通过构建一种可将不同时序长度案件数据匹配为自适应长度数据的方法,并组合向量映射后的天气、作案时间、地理位置等属性,构造多维度特征融合的输入向量;其次,采用自注意力机制生成多视角特征动态融合的向量;最后,通过采用多尺度窗口CNN对多视角特征动态融合向量进行编码后送入分类器,预测出每个地图栅格内的发案态势。在某市盗窃数据集上验证,本文方法在3种地理栅格尺度下,预测准确率最高可达到0.899,显著优于其他对比模型。  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1673-4785  
入藏号:  
CSCD:7352655  
Indexed Date:  
2023-03-24  
Institution Address:  
北京警察学院公安管理系;;北京警察学院北京市公安局警察学院警务情报与数据智能标准实验室, ;;, ;;, 北京;;北京 102202;;102202, 中国  
Institution Address:  
中国传媒大学, 媒体融合与传播国家重点实验室, 北京 100024, 中国  
Institution Address:  
Department of Public Security Management, Beijing Police College;;Standard Laboratory of Police Data and Intelligence of Beijing Public Security Bureau, Beijing Police College, ;;, ;;, Beijing;;Beijing 102202;;102202  
Institution Address:  
Communication University of China, State Key Laboratory of Media Integration and Communication, Beijing 100024, China  
Institution:  
北京警察学院公安管理系;;北京警察学院北京市公安局警察学院警务情报与数据智能标准实验室  
Institution:  
中国传媒大学  
Institution:  
Department of Public Security Management, Beijing Police College;;Standard Laboratory of Police Data and Intelligence of Beijing Public Security Bureau, Beijing Police College  
Institution:  
Communication University of China

### 331. Deep learning-based photoplethysmography biometric authentication for

摘要: Biometric authentication methods have gained prominence as secure and convenient alternatives to traditional passwords and PINs. In this paper, we propose a novel approach for biometric authentication using photoplethysmography (PPG) signals and deep learning techniques. PPG is a non-invasive method that measures variations in blood volume within microvascular tissue beds, and it is typically used for monitoring heart rate and oxygen saturation. Our research leverages the unique characteristics of PPG signals to develop a robust and continuous user verification system. The primary goal of our study is to explore the feasibility and effectiveness of PPG-based biometric authentication, enabling a seamless and secure means of confirming the identity of individuals. We use a diverse dataset of PPG signals from various individuals, ensuring that it encompasses differences in skin tone, age, and other variables that can influence PPG signal characteristics. The collected data undergoes careful preprocessing, including noise removal, baseline correction, and heartbeat segmentation. For the core of our authentication system, we design and train a multiscale feature fusion deep learning (MFFD) model. This model, utilizing a Convolutional Neural Network (CNN) architecture, takes as input the relevant features extracted from PPG signals and learns to differentiate between individuals based on their unique PPG patterns. In this study, the input is constructed by gradually incorporating various features, beginning with a single PPG signal. In this study, the CNN model was trained independently, followed by the implementation of score fusion techniques. Our evaluation demonstrates the effectiveness of the PPG-based biometric authentication system, achieving high accuracy while addressing key security concerns. We consider false acceptance rate (FAR) and false rejection rate (FRR) to assess the system's performance. The model achieves the Accuracy of 99.5 % on BIDMC, 98.6 % on MIMIC, 99.2 % on CapnoBase dataset.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
1568-4946  
eISSN:  
1872-9681  
入藏号:  
WOS:001216666000001  
Indexed Date:  
2024-05-19  
Institution Address:  
Chongqing Univ, Coll Comp Sci, Chongqing 400044, Peoples R China  
Institution Address:  
Columbia Univ, Dept Comp Sci, New York, NY 10027 USA  
Institution Address:  
Princess Nourah bint Abdulrahman Univ, Coll Comp & Informat Sci, Dept Informat Syst, POB 84428, Riyadh 11671, Saudi Arabia  
Institution Address:  
King Saud Univ, Coll Appl Studies & Community Serv, Dept Comp Sci & Engn, Box 22459, Riyadh 11495, Saudi Arabia  
Institution Address:  
Prince Sattam bin Abdulaziz Univ, Coll Comp Engn & Sci, Dept Comp Sci, Al Kharj 16273, Saudi Arabia  
Institution Address:  
Univ Hradec Kralove, Fac Sci, Dept Appl Cybernet, Hradec Kralove 50003, Czech Republic  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 332. MOCNN: A Multiscale Deep Convolutional Neural Network for ERP-Based

摘要: Event-related potentials (ERPs) reflect neurophysiological changes of the brain in response to external events and their associated underlying complex spatiotemporal feature information is governed by ongoing oscillatory activity within the brain. Deep learning methods have been increasingly adopted for ERP-based brain-computer interfaces (BCIs) due to their excellent feature representation abilities, which allow for deep analysis of oscillatory activity within the brain. Features with higher spatiotemporal frequencies usually represent detailed and localized information, while features with lower spatiotemporal frequencies usually represent global structures. Mining EEG features from multiple spatiotemporal frequencies is conducive to obtaining more discriminative information. A multiscale feature fusion octave convolution neural network (MOCNN) is proposed in this article. MOCNN divides the ERP signals into high-, medium-and low-frequency components corresponding to different resolutions and processes them in different branches. By adding mid-and low-frequency components, the feature information used by MOCNN can be enriched, and the required amount of calculations can be reduced. After successive feature mapping using temporal and spatial convolutions, MOCNN realizes interactive learning among different components through the exchange of feature information among branches. Classification is accomplished by feeding the fused deep spatiotemporal features from various components into a fully connected layer. The results, obtained on two public datasets and a self-collected ERP dataset, show that MOCNN can achieve state-of-the-art ERP classification performance. In this study, the generalized concept of octave convolution is introduced into the field of ERP-BCI research, which allows effective spatiotemporal features to be extracted from multiscale networks through branch width optimization and information interaction at various scales.  
在 BIOSIS Citation Index 中的被引频次:  
7  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
16  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
16  
ISSN:  
2168-2267  
eISSN:  
2168-2275  
入藏号:  
WOS:001218673800001  
PubMed ID:  
38713574  
Indexed Date:  
2024-05-21  
Institution Address:  
East China Univ Sci & Technol, Minist Educ, Key Lab Smart Mfg Energy Chem Proc, Shanghai 200237, Peoples R China  
Institution Address:  
East China Univ Sci & Technol, Shenzhen Res Inst, Shenzhen 518063, Peoples R China  
Institution Address:  
Univ Essex, Sch Comp Sci & Elect Engn, Brain Comp Interfacing & Neural Engn Lab, Colchester CO4 3SQ, England  
Institution Address:  
Polish Acad Sci, Syst Res Inst, PL-01447 Warsaw, Poland  
Institution Address:  
RIKEN Adv Intelligence Project, Tensor Learning Lab, Tokyo 1030027, Japan  
Institution Address:  
Tokyo Univ Agr & Technol, Global Innovat Res Inst, Tokyo 1848588, Japan  
Institution:  
Institution:  
East China Univ Sci & Technol  
Institution:  
Institution:  
Institution:  
RIKEN Adv Intelligence Project  
Institution:

### 333. Multiscale Feature Fusion Convolutional Neural Network for Surface

摘要: The detection of surface damage is an important part of the process before remanufacturing a retired steel shaft (RSS). Traditional damage detection is mainly done manually, which is time-consuming and error-prone. In recent years, computer vision methods have been introduced into the community of surface damage detection. However, some advanced typical object detection methods perform poorly in the detection of surface damage on RSS due to the complex surface background and rich diversity of damage patterns and scales. To address these issues, we propose a Faster R-CNN-based surface damage detection method for RSS. To improve the adaptability of the network, we endow it with a feature pyramid network (FPN) as well as adaptable multiscale information modifications to the region proposal network (RPN). In this paper, a detailed study of an FPN-based feature extraction network and the multiscale object detection network is conducted. Experimental results show that our method improves the mean average precision (mAP) score by 8.9% compared with the original Faster R-CNN for surface damage detection of RSS, and the average detection accuracy for small objects is improved by 18.2%. Compared with the current advanced object detection methods, our method is more advantageous for the detection of multiscale objects.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1530-9827  
eISSN:  
1944-7078  
入藏号:  
WOS:001180415700006  
Indexed Date:  
2024-04-02  
Institution Address:  
Dalian Univ Technol, Sch Mech Engn, Dalian 116024, Liaoning, Peoples R China  
Institution Address:  
Dalian Univ Technol, State Key Lab High Performance Precis Mfg, Dalian 116024, Liaoning, Peoples R China  
Institution Address:  
Dalian Univ Technol, Ningbo Inst, Ningbo 315000, Zhejiang, Peoples R China  
Institution Address:  
Army Acad Armored Forces, Natl Key Lab Remfg, Beijing 100072, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Army Acad Armored Forces

### 334. A Road Crack Segmentation Method Based on Transformer and Multi-Scale

摘要: To ensure the safety of vehicle travel, the maintenance of road infrastructure has become increasingly critical, with efficient and accurate detection techniques for road cracks emerging as a key research focus in the industry. The development of deep learning technologies has shown tremendous potential in improving the efficiency of road crack detection. While convolutional neural networks have proven effective in most semantic segmentation tasks, overcoming their limitations in road crack segmentation remains a challenge. To address this, this paper proposes a novel road crack segmentation network that leverages the powerful spatial feature modeling capabilities of Swin Transformer and the Encoder-Decoder architecture of DeepLabv3+. Additionally, the incorporation of a multi-scale coding module and attention mechanism enhances the network's ability to densely fuse multi-scale features and expand the receptive field, thereby improving the integration of information from feature maps. Performance comparisons with current mainstream semantic segmentation models on crack datasets demonstrate that the proposed model achieves the best results, with an MIoU of 81.06%, Precision of 79.95%, and F1-score of 77.56%. The experimental results further highlight the model's superior ability in identifying complex and irregular cracks and extracting contours, providing guidance for future applications in this field.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
eISSN:  
2079-9292  
入藏号:  
WOS:001256447400001  
Indexed Date:  
2024-07-04  
Institution Address:  
Kunming Univ Sci & Technol, Fac Land Resources Engn, Kunming 650093, Peoples R China  
Institution Address:  
Kunming Univ Sci & Technol, City Coll, Dept Earth Sci & Technol, Kunming 650233, Peoples R China  
Institution Address:  
Yunnan Geol Engn Second Invest Inst, Kunming 650093, Peoples R China  
Institution:  
Institution:  
Institution:  
Yunnan Geol Engn Second Invest Inst

### 335. Small Object Detection in UAV Images Based on YOLOv8n

摘要: With the rapid development of unmanned aerial vehicle (UAV) technology, there is an urgent need for high-performance aerial object detection algorithms that are tailored for deployment on drones with limited computing capabilities. This paper proposes a series of improvements to the state-of-the-art YOLOv8 object detector to enhance its detection accuracy and speed for small, partially occluded objects in complex environments. Specifically, we introduce multi-scale feature fusion through additional detection layers, employ conditionally parameterized convolutions to increase representational capacity, and import a dynamic non-monotonic loss function named Wise-IoU to enable more effective regression of bounding boxes. Experiments conducted on the large-scale UAV benchmark dataset VisDrone demonstrate that the improved model achieves state-of-the-art accuracy of 37.6% mAP with 3 M parameters, outperforming other lightweight YOLO detectors and two-stage detectors like Faster R-CNN. The improved model also reaches 40 FPS on an embedded edge device, validating its efficiency and suitability for real-time UAV applications. Through comprehensive quantitative experiments and visual results, this work provides valuable insights and techniques to tailor object detection algorithms for robust and efficient deployment on UAVs with limited onboard computing power.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
6  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
ISSN:  
1875-6891  
eISSN:  
1875-6883  
入藏号:  
WOS:001298794900002  
Indexed Date:  
2024-09-18  
Institution Address:  
Hubei Univ Automot Technol, Sch Elect & Informat Engn, Shiyan 442002, Peoples R China  
Institution Address:  
Dongfeng Automobile Power Parts Co Ltd, Shiyan 442002, Peoples R China  
Institution Address:  
Dongfeng Machine Tool Plant Co Ltd, Shiyan 442002, Peoples R China  
Institution Address:  
Hubei Univ Automot Technol, Sch Econ & Management, Shiyan 442002, Peoples R China  
Institution:  
Institution:  
Dongfeng Automobile Power Parts Co Ltd  
Institution:  
Dongfeng Machine Tool Plant Co Ltd  
Institution:

### 336. Using HVS Dual-Pathway and Contrast Sensitivity to Blindly Assess Image

摘要: Blind image quality assessment (BIQA) aims to evaluate image quality in a way that closely matches human perception. To achieve this goal, the strengths of deep learning and the characteristics of the human visual system (HVS) can be combined. In this paper, inspired by the ventral pathway and the dorsal pathway of the HVS, a dual-pathway convolutional neural network is proposed for BIQA tasks. The proposed method consists of two pathways: the "what" pathway, which mimics the ventral pathway of the HVS to extract the content features of distorted images, and the "where" pathway, which mimics the dorsal pathway of the HVS to extract the global shape features of distorted images. Then, the features from the two pathways are fused and mapped to an image quality score. Additionally, gradient images weighted by contrast sensitivity are used as the input to the "where" pathway, allowing it to extract global shape features that are more sensitive to human perception. Moreover, a dual-pathway multi-scale feature fusion module is designed to fuse the multi-scale features of the two pathways, enabling the model to capture both global features and local details, thus improving the overall performance of the model. Experiments conducted on six databases show that the proposed method achieves state-of-the-art performance.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
1424-8220  
入藏号:  
WOS:000996873700001  
PubMed ID:  
37430884  
Indexed Date:  
2023-06-09  
Institution Address:  
Shenzhen Univ, Dept Artificial Intelligence, Shenzhen 518060, Peoples R China  
Institution Address:  
Educ Univ Hong Kong, Dept Math & Informat Technol, Hong Kong, Peoples R China  
Institution Address:  
Univ Massachusetts Lowell, Dept Elect & Comp Engn, Lowell, MA 01854 USA  
Institution:  
Institution:  
Institution:

### 337. Approximately decoupled component supervision for salient object

摘要: Salient object detection (SOD) aims to find the most attractive object(s) in a scene. In recent years, SOD methods based on deep learning have become the mainstream. Existing methods mostly aggregate the multi-level features extracted by convolutional neural network (CNN) to model the object, or refine the boundary details of the object through multi-scale feature fusion. In this paper, the paradigm of approximately decoupled component supervision is proposed for SOD. Our insight is that the attractive performance of SOD requires explicit modeling of the body and edge of the object with different supervisions. Specifically, we first capture image features through foreground attention (FA) mechanism, cross-block semantic correlation aggregation (CBSC), and resolution-based feature integration (RFI) to make object parts more consistent and complete. Then the detailed edge is obtained by subtracting the body part from the complete mask. By explicitly sampling the body and edge pixels of the salient object, we further optimize the resulting body features and the residual edge features under the supervision of approximate decoupling. Benefiting from the abundant edge information and accurate location information, the framework with various backbone proposed by us can achieve better internal consistency and accurate boundaries of the object. Experimental results on six widely used benchmark datasets demonstrate the superiority and competitiveness of our approach in terms of four popular evaluation metrics. Moreover, the proposed method is an end-to-end saliency detection network without any pre-processing or post-processing.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000770735100001  
Indexed Date:  
2022-03-27  
Institution Address:  
Jilin Univ, Coll Comp Sci & Technol, Changchun 130012, Peoples R China  
Institution Address:  
Jilin Univ, Minist Educ, Key Lab Symbol Computat & Knowledge Engn, Changchun 130012, Peoples R China  
Institution:  
Institution:

### 338. Multi-Fusion Approach for Wood Microscopic Images Identification Based

摘要: With the wide increase in global forestry resources trade, the demand for wood is increasing day by day, especially rare wood. Finding a computer-based method that can identify wood species has strong practical value and very important significance for regulating the wood trade market and protecting the interests of all parties, which is one of the important problems to be solved by the wood industry. This article firstly studies the establishment of wood microscopic images dataset through a combination of traditional image amplification technology and Mix-up technology expansion strategy. Then with the traditional Faster Region-based Convolutional Neural Networks (Faster RCNN) model, the receptive field enhancement Spatial Pyramid Pooling (SPP) module and the multi-scale feature fusion of Feature Pyramid Networks (FPN) module are introduced to construct a microscopic image identification model based on the migration learning fusion model and analyzes the three factors (Mix-up, Enhanced SPP and FPN modules) affecting the wood microscopic image detection model. The experimental results show that the proposed approach can identify 10 kinds of wood microscopic images, and the accuracy rate has increased from 77.8% to 83.8%, which provides convenient conditions for further in-depth study of the microscopic characteristics of wood cells and is of great significance to the field of wood science.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
10  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
eISSN:  
2076-3417  
入藏号:  
WOS:000688616900001  
Indexed Date:  
2021-09-02  
Institution Address:  
Northeast Forestry Univ, Coll Informat & Comp Engn, Harbin 150040, Peoples R China  
Institution Address:  
Northeast Forestry Univ, Coll Mech & Elect Engn, Harbin 150040, Peoples R China  
Institution Address:  
Heilongjiang Forestry Intelligent Equipment Engn, Harbin 150040, Peoples R China  
Institution Address:  
Menoufia Univ, Fac Sci, Dept Math & Comp Sci, Shibin Al Kawm 32511, Egypt  
Institution:  
Institution:  
Institution:  
Heilongjiang Forestry Intelligent Equipment Engn  
Institution:

### 339. Strip Surface Defect Detection Algorithm Based on YOLOv5

摘要: In order to improve the detection accuracy of the surface defect detection of industrial hot rolled strip steel, the advanced technology of deep learning is applied to the surface defect detection of strip steel. In this paper, we propose a framework for strip surface defect detection based on a convolutional neural network (CNN). In particular, we propose a novel multi-scale feature fusion module (ATPF) for integrating multi-scale features and adaptively assigning weights to each feature. This module can extract semantic information at different scales more fully. At the same time, based on this module, we build a deep learning network, CG-Net, that is suitable for strip surface defect detection. The test results showed that it achieved an average accuracy of 75.9 percent (mAP50) in 6.5 giga floating-point operation (GFLOPs) and 105 frames per second (FPS). The detection accuracy improved by 6.3% over the baseline YOLOv5s. Compared with YOLOv5s, the reference quantity and calculation amount were reduced by 67% and 59.5%, respectively. At the same time, we also verify that our model exhibits good generalization performance on the NEU-CLS dataset.  
Web of Science 核心合集中的 "被引频次":  
10  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
11  
eISSN:  
1996-1944  
入藏号:  
WOS:000969944600001  
PubMed ID:  
37049103  
Indexed Date:  
2023-05-05  
Institution Address:  
Guangdong Univ Technol, Sch Mech & Elect Engn, Guangzhou 510006, Peoples R China  
Institution Address:  
Guangdong Acad Sci, Inst Intelligent Mfg, Guangzhou 510070, Peoples R China  
Institution Address:  
Guangdong Prov Key Lab Modern Control Technol, Guangzhou 510070, Peoples R China  
Institution:  
Institution:  
Institution:  
Guangdong Prov Key Lab Modern Control Technol

### 340. Image Classification Method Based on Multi-Scale Convolutional Neural

摘要: Traditional convolutional neural networks (CNNs) typically use fixed scale convolutional kernels for feature extraction when processing image classification tasks, while ignoring the multi-scale information present in the image. To overcome this limitation, we propose an algorithm based on multi-scale CNNs, which capture features at different levels by introducing convolutional kernels of different scales into the convolutional layer. In this study, we first designed a multi-scale convolutional layer consisting of multiple convolutional kernels of different scales to extract multi-scale features of the image. To further enhance classification performance, we introduced a multi-scale feature fusion module that can effectively fuse features of different scales and classify them through a fully connected layer. Then we conducted extensive experiments on several commonly used image classification datasets. The experimental results show that this network can not only effectively identify and locate hyperspectral image targets in different scenarios, but also reduce missed detections and false positives during the detection process. The average accuracy of the improved model has been improved, and the recognition accuracy of some small markers affected by external factors such as occlusion and lighting has also been improved. In addition, by comparing the detection effect of a single image, the progressiveness and anti-leakage ability of the improved model are proved. The image classification method based on multi-scale CNNs has broad application prospects in image recognition and feature extraction, and can provide valuable reference and reference for research in related fields.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0218-1266  
eISSN:  
1793-6454  
入藏号:  
WOS:001193334000002  
Indexed Date:  
2024-04-05  
Institution Address:  
Guizhou Univ Commerce, Sch Comp & Informat Engn, Guiyang 550014, Guizhou, Peoples R China  
Institution:

### 341. U-DPnet: an ultralight convolutional neural network for the detection of

摘要: Efficient and accurate detection of apples is critical for the successful implementation of harvesting robots in orchards. However, due to limited memory resources on robotic platforms, it is imperative to develop lightweight detection algorithms that can operate in real-time. To address this challenge, we propose an ultralight convolutional neural network, U-DPnet, based on depth-separable convolution. Our approach incorporates the cross-stage deep separable module (CDM) and the multi-cascade deep separable module (MDM) in the backbone for nonlinear unit addition and attention mechanisms, which reduce the volume of the network while improving the feature representation capability. A simplified bi-directional feature pyramid network (BiFPN) is constructed in the neck for multi-scale feature fusion, and Adaptive feature propagation (AFP) is designed between the neck and the backbone for smooth feature transitions across different scales. To further reduce the network volume, we develop a uniform channel downsampling and network weight-sharing strategy. Multiple loss functions and label assignment strategies are used to optimize the training process. The performance of U-DPnet is verified on a homemade Apple dataset. Experimental results demonstrate that U-DPnet achieves detection accuracy and speed comparable to that of the 7 SOTA models. Moreover, U-DPnet exhibits an absolute advantage in model volume and computations (only 1.067M Params and 0.563G FLOPs, 39.79% and 36.36% less than yolov5-n).  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1861-8200  
eISSN:  
1861-8219  
入藏号:  
WOS:001012461100001  
Indexed Date:  
2023-07-03  
Institution Address:  
Northwestern Polytech Univ, Sch Automat, Xian 710100, Shaanxi, Peoples R China  
Institution:

### 342. A light-weight stereo matching network based on multi-scale features

摘要: In recent years, convolutional-neural-network based stereo matching methods have achieved significant gains compared to conventional methods in terms of both speed and accuracy. Current state-of-the-art disparity estimation algorithms require many parameters and large amounts of computational resources and are not suited for applications on edge devices. In this paper, an end-to-end light-weight network (LWNet) for fast stereo matching is proposed, which consists of an efficient backbone with multi-scale feature fusion for feature extraction, a 3D U-Net aggregation architecture for disparity computation, and color guidance in a 2D convolutional neural network (CNN) for disparity refinement. MobileNetV2 is adopted as an efficient backbone in feature extraction. The channel attention module is applied to improve the representational capacity of features and multi-resolution information is adaptively incorporated into the cost volume via cross-scale connections. Further, a left-right consistency check and color guidance refinement are introduced and a robust disparity refinement network is designed with skip connections and dilated convolutions to capture global context information and improve disparity estimation accuracy with little computational cost and memory space. Extensive experiments on Scene Flow, KITTI 2015, and KITTI 2012 demonstrate that the proposed LWNet achieves competitive accuracy and speed when compared with state-of-the-art stereo matching methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1751-9659  
eISSN:  
1751-9667  
入藏号:  
WOS:000935722400001  
Indexed Date:  
2023-03-16  
Institution Address:  
Guizhou Univ, Sch Mech Engn, Guiyang, Peoples R China  
Institution Address:  
Guizhou Acad Agr Sci, Guizhou Tea Res Inst, Guiyang, Peoples R China  
Institution Address:  
Peking Univ, Sch Elect & Comp Engn, Shenzhen Grad Sch, Shenzhen, Peoples R China  
Institution Address:  
Guizhou Educ Univ, Sch Math & Big Data, Guiyang, Peoples R China  
Institution Address:  
Guangzhou Med Univ, Qingyuan Peoples Hosp, Affiliated Hosp 6, Sch Biomed Engn, Guangzhou, Peoples R China  
Institution Address:  
Liupanshui Normal Coll, Sch Min & Civil Engn, Liupanshui, Peoples R China  
Institution Address:  
Guizhou Univ, Sch Mech Engn, Guiyang 550025, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 343. MFF-DenseNet: Densely Connected Convolutional Network With Multi-Scale

摘要: Magnetotelluric (MT) is a geophysical technique for detecting subsurface electrical structures. However, MT data collected in areas with frequent human activity often encounter various types of electromagnetic (EM) noise, which can mask or distort the signals we aim to analyze. Over the past decades, data processing methods based on deep learning has become the focus of multiple disciplines. Training neural networks to identify and handle noise has been proven effective in reducing the impact of noise. Therefore, ensuring the neural network accurately learns the noise and signal characteristics during the training is crucial. Against this background, we propose a multi-scale feature fusion technique based on the densely connected network and apply it to processing MT data. First, we construct a data set resembling the noise in field data and use it to train the network. Leveraging dense connections, we extract feature maps of EM noise from noisy data and utilize Spatial Pyramid Pooling to integrate feature maps of various scales, enabling the network to capture features of the noise precisely. At the same time, we reduce the computation of feature fusion by introducing the Channel-wise Squeezed Layer to compress the channels of the feature maps. Ultimately, we apply the trained model to the field noisy data. The results of synthetic and field data demonstrate that our method suppresses low-amplitude and continuous high-amplitude noise while preserving low-frequency valuable signal. Apparent resistivity-phase curves and polarization direction shows a noticeable improvement in the mid and low-frequency bands with our method.  
Magnetotelluric (MT) is a geophysical technique that helps scientists study the electrical properties in the subsurface. This method is helpful for finding minerals and understanding the structure of the Earth. However, MT data collected in mining areas are susceptible to the influence of human activities, which impedes geophysical studies. In recent years, researchers have started using neural networks to clean up noisy MT data. Our research introduces a new neural network called a multi-scale feature fusion densely connected network which is directly applied to the time-series data. Using the multi-scale feature fusion method, our method separates noise and signals more accurately. Consequently, it can reduce the impact of noise in mining areas, leading to clearer and more accurate geophysical studies.  
We propose a multi-scale feature fusion technique based on DenseNet to separate signal and noise from the perspective of feature maps Our method shows more accurate denoising results in synthetic and field data compared to wavelet transform and Convolutional Neural Network methods Our method demonstrates robust performance in new regions by using transfer learning with a small amount of training data  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-9313  
eISSN:  
2169-9356  
入藏号:  
WOS:001315629800001  
Indexed Date:  
2024-09-25  
Institution Address:  
Hunan Normal Univ, Coll Informat Sci & Engn, Hunan Prov Key Lab Intelligent Comp & Language Inf, Changsha, Peoples R China  
Institution Address:  
Cent South Univ, Key Lab Metallogen Predict Nonferrous Met & Geol E, Monitoring Minist Educ, Changsha, Peoples R China  
Institution Address:  
Minist Nat Resources, Tech Innovat Ctr Coverage Area Deep Resources Expl, Hefei, Peoples R China  
Institution:  
Institution:  
Institution:

### 344. DSWMamba: A deep feature fusion mamba network for detection of asphalt

摘要: The maintenance of asphalt pavements is vital for national infrastructure. Timely detection of pavement distress ensures optimal road performance and extends pavement lifespan. However, the various distress types and complex detection environments pose significant challenges. To overcome these, we propose a novel framework named DSWMamba, which is the first application of Mamba in asphalt pavement distress detection. This model incorporates targeted optimizations via the Selective Scanning Modules (SSM) to meet the unique demands of distress detection. To overcome the limitations of the SSM in receptive fields and poor image localization, we develop the Depth Fusion Selective Scan Block (DFSS), which integrates a Fusion Perception Block (FP-Block). This design expands the receptive field and strengthens local feature extraction. Additionally, the Vision Shuffle Down module(VSD) preserves both multiscale feature information, further enhancing feature representation. Unlike Transformer-based models, DSWMamba leverages a Separate Multidimensional Attention (SMD-Attention) mechanism for multi-path fusion, enabling global modeling. To improve the detection of large-scale features, we develop a Dynamic Separable align Head (DSA-Head), which decouples classification and localization tasks, significantly boosting feature recognition accuracy. Extensive experiments demonstrate that DSWMamba surpasses 18 target detection models, including CNN, Transformer frameworks and the latest YOLO and RTDETR series, on three datasets. DSWMamba improves transverse crack detection accuracy on the Asphalt Pavement Distress Dataset by 13.1% compared to Mamba YOLO while reducing parameters by 34.6% and computational cost by 44.6%. We release the Asphalt Pavement Distress Dataset, comprising 13,129 images of six distress types, offering a valuable resource for future research.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0950-0618  
eISSN:  
1879-0526  
入藏号:  
WOS:001444195500001  
Indexed Date:  
2025-03-21  
Institution Address:  
Guangxi Univ, Sch Comp Elect & Informat, 100 Daxue Rd, Nanning 530004, Guangxi, Peoples R China  
Institution Address:  
Guangxi Univ, Sch Elect Engn, Nanning 530004, Peoples R China  
Institution Address:  
Guangxi Univ, Sch Civil Engn & Architecture, Nanning 530004, Peoples R China  
Institution:  
Institution:  
Institution:

### 345. Adaptive multi-scale attention convolution neural network for

摘要: This paper proposes a novel approach named adaptive multi-scale attention convolution neural network (AmaCNN) to accurately detect cross-domain faults with very few labelled data. In AmaCNN, multi-scale feature fusion CNN (MSFFCNN) with a multi-level attention scheme (MLAS) extracts multi-scale less-noise features from source and target domains. Considering the domain shift and semantic difference in the two domain features, a cross-domain adaption (CDA) scheme is applied. Significantly, the extracted domain features are measured with correlation alignment (CORAL) distance to minimize the domain shift first. Then, semantic alignment (SA) loss aligns and separates domain-invariant features point-by-point. Therefore, the proposed AmaCNN could learn rich multi-scale, less-noise, domain-invariant, and semantic-alignment features using limited training samples to detect cross-fault accurately. The experimental results on three real data sets confirmed its priority and reliability. Besides, the in-depth analysis has confirmed each component's effectiveness and CDA's good generality.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
29  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
29  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001070592200001  
Indexed Date:  
2023-10-05  
Institution Address:  
Pukyong Natl Univ, Ind Sci Technol Res Ctr, Busan 608737, South Korea  
Institution Address:  
Pukyong Natl Univ, Informat Syst, Busan 608737, South Korea  
Institution:  
Institution:

### 346. SLMS-SSD: Improving the balance of semantic and spatial information in

摘要: With the development of deep learning technology, the research on convolutional neural network-based object detection is becoming more and more mature. However, most methods are unsatisfactory in dealing with the issue of semantic and spatial information imbalance. In this article, we extend the single-shot multibox detector SSD and propose a self-learning multi-scale object detection network by balancing the semantic information and spatial information, named SLMS-SSD. We first construct a shallow feature enhancement module to enhance the representation of small objects by extracting richer context information. Second, in terms of feature connectivity, we design a multi-scale feature selection module for intermediate layer features with a combination of top-down and direct up-samplings. Finally, in terms of feature strength, we design a self-learning feature fusion module for measuring the feature importance. We validate our model on the PASCAL VOC and MS COCO datasets, and the results demonstrate that it can effectively improve the accuracy of object detection, especially the accuracy of small object detection.  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
11  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
12  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:000832966000001  
Indexed Date:  
2022-08-11  
Institution Address:  
Beijing Univ Chem Technol, Coll Informat Sci & Technol, Beijing 100029, Peoples R China  
Institution Address:  
Univ Sci & Technol China, Dept Automat, Hefei 230027, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Management & Control Complex Syst, Beijing 100190, Peoples R China  
Institution:  
Institution:  
Institution:

### 347. Change Detection of High-Resolution Remote Sensing Images Through

摘要: One of the major challenges in the change detection (CD) of high-resolution remote sensing images is the high requirement for computational resources. Besides, to get the best change detection result, it must spot only the important changes while omitting unimportant ones, which requires learning complex interactions between multi-scale objects on the images. Despite Convolution Neural Network (CNN) efficiently extracting features from such images, it has a limited receptive field resulting in sub-optimal representation. On the other hand, Vision Transformer (ViT) can capture long-range dependencies. Still, it suffers from quadratic complexity concerning the number of image patches, especially for high-resolution images. Furthermore, both approach can not model the interactions among multi-scale image patches, which is essential for a model to fully understand the natural images. We propose FocalCD, a CD method based on a recently proposed focal modulation architecture capable of learning short and long dependencies to solve this problem. It is attention-free and does not suffer from quadratic complexity. Also, it supports learning multi-scale interaction by adaptively selecting the discriminator regions from multi-scale levels. Besides the efficient yet powerful encoder, FocalCD has an effective multi-scale feature fusion and pyramidal decoder network. FocalCD achieves strong empirical results on various CD datasets, including CDD, LEVIR-CD, and WHU-CD. It reaches F1 scores of 0.9851, 0.952, and 0.9616 on datasets CDD, LEVIR-CD, and WHU-CD outperforming state-of-the-art CD methods while having comparable or even lower computation complexity.  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
2169-3536  
入藏号:  
WOS:001030576500001  
Indexed Date:  
2023-07-31  
Institution Address:  
Univ Indonesia, Fac Comp Sci, Depok City 16424, Indonesia  
Institution:

### 348. Water Body Extraction in Remote Sensing Imagery Using Domain

摘要: A water body is a common object in remote sensing images and high-quality water body extraction is important for some further applications. With the development of deep learning (DL) in recent years, semantic segmentation technology based on deep convolution neural network (DCNN) brings a new way for automatic and high-quality body extraction from remote sensing images. Although several methods have been proposed, there exist two major problems in water body extraction, especially for high resolution remote sensing images. One is that it is difficult to effectively detect both large and small water bodies simultaneously and accurately predict the edge position of water bodies with DCNN-based methods, and the other is that DL methods need a large number of labeled samples which are often insufficient in practical application. In this paper, a novel SFnet-DA network based on the domain adaptation (DA) embedding selective self-attention (SSA) mechanism and multi-scale feature fusion (MFF) module is proposed to deal with these problems. Specially, the SSA mechanism is used to increase or decrease the space detail and semantic information, respectively, in the bottom-up branches of the network by selective feature enhancement, thus it can improve the detection capability of water bodies with drastic scale change and can prevent the prediction from being affected by other factors, such as roads and green algae. Furthermore, the MFF module is used to accurately acquire edge information by changing the number of the channel of advanced feature branches with a unique fusion method. To skip the labeling work, SFnet-DA reduces the difference in feature distribution between labeled and unlabeled datasets by building an adversarial relationship between the feature extractor and the domain classifier, so that the trained parameters of the labeled datasets can be directly used to predict the unlabeled images. Experimental results demonstrate that the proposed SFnet-DA has better performance on water body segmentation than state-of-the-art methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
11  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
11  
eISSN:  
2072-4292  
入藏号:  
WOS:000839826800001  
Indexed Date:  
2022-08-22  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Astronaut, Nanjing 211106, Peoples R China  
Institution:

### 349. A Study on Maize Leaf Pest and Disease Detection Model Based on

摘要: The detection and accurate positioning of agricultural pests and diseases can significantly improve the effectiveness of disease and pest control and reduce the cost of prevention and control, which has become an urgent need for crop production. Aiming at the low precision of maize leaf pest and disease detection, a new model of maize leaf pest and disease detection using attention mechanism and multi-scale features was proposed. Our model combines a convolutional block attention module (CBAM) with the ResNet50 backbone network to suppress complex background interference and enhance feature expression in specific regions of the maize leaf images. We also design a multi-scale feature fusion module that aggregates local and global information at different scales, improving the detection performance for objects of varying sizes. This module reduces the number of parameters and enhances efficiency by using a lightweight module and replacing the deconvolutional layer. Experimental results on a natural environment dataset demonstrate that our proposed model achieves an average detection accuracy of 85.13%, which is 9.59% higher than the original CenterNet model. The model has 24.296 M parameters and a detection speed of 23.69 f/s. Compared with other popular models such as SSD-VGG, YOLOv5, Faster-RCNN, and Efficientdet-D0, our proposed model demonstrates superior performance in the fast and accurate detection of maize leaf pests and diseases. This model has practical applications in the identification and treatment of maize pests and diseases in the field, and it can provide technical support for precision pesticide application. The trained model can be deployed to a web client for user convenience.  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
eISSN:  
2076-3417  
入藏号:  
WOS:001071077700001  
Indexed Date:  
2023-10-08  
Institution Address:  
Shaanxi Univ Sci & Technol, Sch Elect & Control Engn, Xian 710021, Peoples R China  
Institution:

### 350. Crowd Counting Method Based on Multiscale Perception and Image

摘要: Crowd counting is widely used in public security,video surveillance and other fields,but due to factors such as target occlusion,background interference,and crowd size changes,the accuracy of crowd counting models is reduced. Based on a deep learning convolutional neural network architecture,a crowd counting method according to multi-scale perception and image association is proposed. The multi-scale perception model includes a primary feature extraction network,a multi-scale feature extraction module,a feature fusion module and a back-end architecture to extract the multi-scale features of the image,so as to adapt to the change of scale,while the image correlation model uses the feature correlation module and the fusion module to associate input images with coherent images,and improve the quality of predicted density maps through learning deep correlations between images. The experimental results on public datasets such as ShanghaiTech Part\_A,Part\_B and UCF\_CC\_50 show that the proposed method has good performance on three indicators of MAE,RMSE and SSIM.  
摘要:  
人群计数广泛应用于公共安防、视频监控等领域,但由于目标遮挡、背景干扰以及人群尺度变化等因素的影响,人群计数模型的准确率有所降低。基于深度学习卷积神经网络架构,提出了一种基于多尺度感知和图像关联的人群计数方法。其中,多尺度感知模型包括初级特征提取网络、多尺度特征提取模块、特征融合模块和一个后段架构用来提取图像的多尺度特征,从而适应尺度的变化;而图像关联模型使用特征关联模块和融合模块将输入图像与相干图像进行联系,通过学习图像之间的深层关联性来提升预测密度图的质量。在ShanghaiTech Part\_A、Part\_B和UCF\_CC\_50等公开数据集上的实验结果表明,提出的方法在MAE、RMSE和SSIM三项指标上均有较好性能。  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1004-1699  
入藏号:  
CSCD:7617336  
Indexed Date:  
2024-02-08  
Institution Address:  
浙江工业大学信息工程学院, 杭州, 浙江 310012, 中国  
Institution Address:  
College of Information Engineering, Zhejiang University of Technology, Hangzhou, Zhejiang 310012, China  
Institution:  
浙江工业大学信息工程学院  
Institution:  
College of Information Engineering, Zhejiang University of Technology

### 351. Integrating Multi-Scale Feature Boundary Module and Feature Fusion With

摘要: The skin, a crucial organ, plays a protective role in the human body, emphasizing the significance of early detection of skin diseases to prevent potential progression to skin cancer. The challenge lies in diagnosing these diseases at their early stages, where visual resemblance complicates differentiation, highlighting the need for an innovative automated method for precisely identifying skin lesions in biomedical images. This paper introduces a holistic methodology that combines DenseNet, multi-scale feature boundary module (MFBM), and feature fusion and decoding engine (FFDE) to tackle challenges in existing deep-learning image segmentation methods. Furthermore, a convolutional neural network model is designed for the classification of segmented images. The DenseNet encoder efficiently extracts features at four resolution levels, leveraging dense connectivity to capture intricate hierarchical features. The proposed MFBM plays a crucial role in extracting boundary information, employing parallel dilated convolutions with various dilation rates for effective multi-scale information capture. To overcome potential disadvantages related to the conversion of features during segmentation, our approach ensures the preservation of context features. The proposed FFDE method adaptively fuses features from different levels, restoring skin lesion location information while preserving local details. The evaluation of the model is conducted on the HAM10000 dataset, which consists of 10 015 dermoscopy images, yielding promising results.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0899-9457  
eISSN:  
1098-1098  
入藏号:  
WOS:001307467400001  
Indexed Date:  
2024-09-13  
Institution Address:  
Francis Xavier Engn Coll, Dept Comp Sci & Engn, Tirunelveli, Tamil Nadu, India  
Institution:  
Francis Xavier Engn Coll

### 352. Semantic segmentation method for remote sensing images based on

摘要: To improve the accuracy of semantic segmentation for remote sensing images and address the loss problem of small-sized target information during feature extraction by Deep Convolutional Neural Network (DCNN), a semantic segmentation method based on multi-scale feature fusion named FuseSwin was proposed. Firstly, an Attention Enhancement Module (AEM)was introduced in the Swin Transformer to highlight the target area and suppress background noise. Secondly, the Feature Pyramid Network (FPN)was used to fuse the detailed information and high-level semantic information of the multi-scale features to complement the features of the target. Finally, the Atrous Spatial Pyramid Pooling (ASPP)module was used to capture the contextual information of the target from the fused feature map and further improve the model segmentation accuracy. Experimental results demonstrate that the proposed method outperforms current mainstream segmentation methods. The mean Pixel Accuracy (mPA)and mean Intersection over Union (mIoU)of the proposed method on Potsdam remote sensing dataset are 2.34 and 3.23 percentage points higher than those of DeepLabV3 method, and 1.28 and 1.75 percentage points higher than those of SegFormer method. Additionally, the proposed method was applied to identify and segment oyster rafts in high-resolution remote sensing images of the Maowei Sea in Qinzhou, Guangxi, and achieved Pixel Accuracy(PA)and Intersection over Union(IoU)of 96.21% and 91.70%, respectively.  
摘要:  
为提高遥感图像语义分割精度,解决深度卷积神经网络(DCNN)特征提取过程中小尺寸目标信息丢失的问题,提出一种基于多尺度特征融合的语义分割方法FuseSwin。首先,在Swin Transformer中引入注意力增强模块(AEM),以突出目标所在区域并抑制背景噪声的干扰;其次,利用特征金字塔网络(FPN)融合多尺度特征的细节信息和高级语义信息,以补充目标的特征;最后,通过空洞空间金字塔池化(ASPP)模块从融合特征图中进一步捕获目标的上下文信息,提升模型分割精度。实验结果表明,所提方法在Potsdam遥感数据集上的平均像素准确率(mPA)和平均交并比(mIoU),与DeepLabV3方法相比,分别提高了2.34、3.23个百分点;与SegFormer方法相比,分别提高了1.28、 1.75个百分点,优于目前主流的分割方法。此外,将所提方法实际应用于广西钦州茅尾海的高分辨率遥感图像中的蚝排识别与分割,分别取得96.21%、91.70%的像素准确率(PA)和交并比(IoU)。  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1001-9081  
入藏号:  
CSCD:7686811  
Indexed Date:  
2024-06-14  
Institution Address:  
广西大学计算机与电子信息学院;;北部湾大学, ;;广西海洋工程装备与技术重点实验室, 南宁;;钦州, ;;广西 530004;;535011  
Institution Address:  
广西大学计算机与电子信息学院, 南宁, 广西 530004, 中国  
Institution Address:  
广西大学计算机与电子信息学院;;广西大学, ;;广西多媒体通信与网络技术重点实验室, 南宁;;南宁, ;; 530004;;530004  
Institution Address:  
School of Computer, Electronics and Information, Guangxi University;;Beibu Gulf University, ;;Guangxi Key Laboratory of Marine Engineering Equipment and Technology, Nanning;;Qinzhou, Guangxi;;Guangxi 530004;;535011  
Institution Address:  
School of Computer, Electronics and Information, Guangxi University, Nanning, Guangxi 530004, China  
Institution Address:  
School of Computer, Electronics and Information, Guangxi University;;Guangxi University, ;;Guangxi Key Laboratory of Multimedia Communications and Network Technology, Nanning;;Nanning, Guangxi;;Guangxi 530004;;530004  
Institution:  
广西大学计算机与电子信息学院;;北部湾大学  
Institution:  
广西大学计算机与电子信息学院  
Institution:  
广西大学计算机与电子信息学院;;广西大学  
Institution:  
School of Computer, Electronics and Information, Guangxi University;;Beibu Gulf University  
Institution:  
School of Computer, Electronics and Information, Guangxi University  
Institution:  
School of Computer, Electronics and Information, Guangxi University;;Guangxi University

### 353. PASTFNet: a paralleled attention spatio-temporal fusion network for

摘要: Micro-expressions (MEs) play such an important role in predicting a person's genuine emotions, as to make micro-expression recognition such an important resea rch focus in recent years. Most recent researchers have made efforts to recognize MEs with spatial and temporal information of video clips. However, because of their short duration and subtle intensity, capturing spatio-temporal features of micro-expressions remains challenging. To effectively promote the recognition performance, this paper presents a novel paralleled dual-branch attention-based spatio-temporal fusion network (PASTFNet). We jointly extract short- and long-range spatial relationships in spatial branch. Inspired by the composite architecture of the convolutional neural network (CNN) and long short-term memory (LSTM) for temporal modeling, we propose a novel attention-based multi-scale feature fusion network (AMFNet) to encode features of sequential frames, which can learn more expressive facial-detailed features for it implements the integrated use of attention and multi-scale feature fusion, then design an aggregation block to aggregate and acquire temporal features. At last, the features learned by the above two branches are fused to accomplish expression recognition with outstanding effect. Experiments on two MER datasets (CASMEII and SAMM) show that the PASTFNet model achieves promising ME recognition performance compared with other methods.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0140-0118  
eISSN:  
1741-0444  
入藏号:  
WOS:001171732000001  
PubMed ID:  
38413518  
Indexed Date:  
2024-03-08  
Institution Address:  
Xinjiang Univ, Sch Informat Sci & Engn, Urumqi, Peoples R China  
Institution Address:  
Xinjiang Univ, Sch Software, Urumqi, Peoples R China  
Institution Address:  
Key Lab Signal Detect & Proc Xinjiang Uygur Autono, Urumqi 830014, Peoples R China  
Institution:  
Institution:  
Institution:  
Key Lab Signal Detect & Proc Xinjiang Uygur Autono

### 354. Image classification method for tomato leaf deficient nutrient elements

摘要: This study aims to realize the accurate identification of deficient nutrients elements in tomato leaves. An experiment was conducted on the lack of nutrients in the climate chamber in the laboratory of Big Data Intelligence Department of the Beijing Academy of Agriculture and Forestry Sciences, China. An artificial climate chamber was also selected to regulate the growth environment factors of tomato plants for the specific lack of nutrients. Three types of nutrient deficiency groups were set, namely nitrogen deficiency, phosphorus deficiency, and potassium deficiency, as well as a normal control group. The experiment was started with the appearance of nutrient-deficient traits in seedlings, and the images of nutrient-deficient leaves were then collected according to the growth stages. The experimental results show that there were diversity and differences in the traits of tomato nutrient deficiency. Specifically, there were relatively small changes of leaves in the early stage of tomato nutrient deficiency. Furthermore, it was difficult to capture the details and textures, due to the smaller area of traits. For example, the manifestation of phosphorus deficiency was that the leaves gradually turn purple along the veins. The trait details were hardly identified in the early stage of phosphorus deficiency, due mainly to the mostly small vein structure. Particularly, tomato leaves under different conditions of nutrient deficiency presented similar color and texture characteristics at a certain stage. For example, the leaves were both slightly yellow in the early stage of nitrogen deficiency and the early stage of potassium deficiency. The only slight difference was the characteristic display of morphology in the size of characteristic areas. There were obvious differences in color and texture at different stages under the same nutrient deficiency. The images were collected from the climate chamber to serve as the experimental data. An attempt was made on the inconsistency of feature area size, and the difficulty of feature extraction, resulting from the different types of nutrient deficiency, the insignificant early traits of nutrient deficiency, and the large differences in the characteristics of each growth period. Therefore, an image classification was proposed for the nutrient deficiency of tomato leaves using an attention mechanism and multi-scale feature fusion convolutional neural network (MSFF & AM-CNNs). First of all, a multi-scale feature fusion (MSFF) module was set for nutrient deficiency traits, due to the low efficiency of a fixed-scale convolution kernel for different sizes. The MSFF input image was carried out with multi-channel feature stitching after the MSFF convolution kernel of multiple scales, where the shallow image was multiplied while expanding the number of channels. As such, the fusion of scale features was adopted in this structure. Secondly, an MSFF&AM module was used to improve the large-scale convolutional layer for the extraction of shallow features using the attention mechanism (CBAM). A multi-scale fusion of Bottleneck was also utilized to improve the Dense Block for the extraction of deep features. Deep-MSFF Block aimed to combine the attention mechanism and the MSFF module, where the multiple feature channels were selectively emphasized the global multi-scale information feature function. The recalibration of features in nitrogen deficiency was improved on the tomato leaves the classification accuracy. Finally, a Focal Loss function was introduced as the loss function to reduce the weight of easy-to-differentiate samples.  
摘要:  
针对番茄早期缺素性状不明显及各生长期特征差异较大所导致的特征区域尺寸不一致、难提取、难辩别等问题,提出了一种基于注意力机制及多尺度特征融合卷积神经网络(Multi-Scale Feature Fusion Convolutional Neural Networks based on Attention Mechanism,MSFF-AM-CNNs)的番茄叶片缺素图像分类方法。首先根据番茄叶片缺素特点提出了多尺度特征融合结构(Multi-Scale Feature Fusion Module,MSFF Module);其次在DenseNet基础上,结合浅层网络主要提取纹理、细节特征,深层网络主要提取轮廓、形状特征的特点分别提出具有针对性的特征提取方法,通过不同形式引入注意力机制及多尺度特征融合结构,使全局多尺度信息融合多个特征通道、选择性地强调信息特征并达到对特征精准定位的功能;同时引入Focal Loss函数以减少易分类样本的权重。试验结果表明,MSFF-AM-CNNs的平均召回率、平均F1得分、平均准确率较原模型DenseNet-121均大幅提升,其中缺氮和缺钾叶片的准确率分别提高了8.06和6.14个百分点,召回率分别提高了6.31和5.00个百分点,F1得分分别提高了7.25和5.55个百分点,平均识别准确率可达95.92%,具有较高的识别准确率及广泛的适用性,能够满足番茄叶片缺素图像的高精度分类需求,可为植物叶片缺素识别提供参考。  
在中国科学引文数据库中的被引频次:  
12  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
14  
ISSN:  
1002-6819  
入藏号:  
CSCD:7090950  
Indexed Date:  
2022-01-21  
Institution Address:  
西北农林科技大学信息工程学院;;国家农业信息化工程技术研究中心;;北京市农业信息技术研究中心, ;;国家农业信息化工程技术研究中心;;, 杨凌;;;;, ;;北京;;北京 712100;;100097;;100097  
Institution Address:  
国家农业信息化工程技术研究中心;;北京市农业信息技术研究中心, 国家农业信息化工程技术研究中心;;, ;;, 北京;;北京 100097;;100097, 中国  
Institution Address:  
College of Information Engineering, Northwest A&F University;;National Engineering Research Center for Information Technology In Agriculture;;Beijing Research Center for Information Technology In Agriculture, ;;National Engineering Research Center for Information Technology In Agriculture;;, Yangling;;;;, ;;Beijing;;Beijing 712100;;100097;;100097  
Institution Address:  
National Engineering Research Center for Information Technology In Agriculture;;Beijing Research Center for Information Technology In Agriculture, National Engineering Research Center for Information Technology In Agriculture;;, ;;, Beijing;;Beijing 100097;;100097  
Institution:  
西北农林科技大学信息工程学院;;国家农业信息化工程技术研究中心;;北京市农业信息技术研究中心  
Institution:  
国家农业信息化工程技术研究中心;;北京市农业信息技术研究中心  
Institution:  
College of Information Engineering, Northwest A&F University;;National Engineering Research Center for Information Technology In Agriculture;;Beijing Research Center for Information Technology In Agriculture  
Institution:  
National Engineering Research Center for Information Technology In Agriculture;;Beijing Research Center for Information Technology In Agriculture

### 355. A Gesture Recognition Method Based on Recurrent Spatiotemporal Depth

摘要: To solve the problem of weak robustness and low precision of existing hand gesture recognition models induced by lack of spatiotemporal information,a hand gesture recognition model based on recurrent spatial and temporal deep neural network is proposed to improve the characterization ability for surface EMG(sEMG)signals.Firstly,a multi-channel convolutional neural network is designed and integrated into the bidirectional recurrent neural network to extract the spatiotemporal characteristics information with strong discrimination.Secondly,channel attention mechanism is used to catch the channel importance information in spatiotemporal characteristics,then an attention module based on spatiotemporal characteristics is designed to further enhance the spatiotemporal characteristics information.Thirdly,based on the ideology of feature pyramid network,a multi-scale feature fusion module is designed to acquire multistage feature information based on multi-scale and multi-angle aspects to improve the decoding ability of the model to electromyography signals.Finally,the proposed hand gesture recognition model is tested based on a large hand gesture recognition database of Ninapro.The results show that the representation capability for sEMg signals is effectively improved by the proposed method.It provides reference for the deep learning modeling work of human body hand gesture recognition.  
摘要:  
针对表面肌电信号解码模型因缺乏时空信息等重要性表征,面临解码精度低、鲁棒性差等问题,提出了一种基于循环时空深度神经网络的手势识别模型,来提高挖掘表面肌电信号的表征能力。首先,设计多通道卷积神经网络,并融入双向循环神经网络来提取强判别力的时空特征信息。其次,采用通道注意力机制来捕捉时空特征中通道重要性信息,设计基于时空特征的注意力模块以进一步增强时空特征信息。同时,基于特征金字塔网络思想来设计多尺度特征融合模块,从多尺度、多角度获取多级特征信息,提高模型对肌电信号的解码能力。最后,将所提出的手势识别模型在大型手势识别数据库Ninapro上进行测试,结果表明所提方法能有效提高对表面肌电信号的表征挖掘能力,为人体手势动作识别的深度学习建模工作提供借鉴意义。  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1004-1699  
入藏号:  
CSCD:7782763  
Indexed Date:  
2024-10-18  
Institution Address:  
浙江工业大学信息工程学院, 杭州, 浙江 310023, 中国  
Institution Address:  
College of Information Engineering,Zhejiang University of Technology, Hangzhou, Zhejiang 310023, China  
Institution:  
浙江工业大学信息工程学院  
Institution:  
College of Information Engineering,Zhejiang University of Technology

### 356. FCSU-Net: A novel full-scale Cross-dimension Self-attention U-Net with

摘要: Recently, ViT and CNNs based on encoder-decoder architecture have become the dominant model in the field of medical image segmentation. However, there are some deficiencies for each of them: (1) It is difficult for CNNs to capture the interaction between two locations with consideration of the longer distance. (2) ViT cannot acquire the interaction of local context information and carries high computational complexity. To optimize the above deficiencies, we propose a new network for medical image segmentation, which is called FCSU-Net. FCSU-Net uses the proposed collaborative fusion of multi-scale feature block that enables the network to obtain more abundant and more accurate features. In addition, FCSU-Net fuses full-scale feature information through the FFF (Full-scale Feature Fusion) structure instead of simple skip connections, and establishes long-range dependencies on multiple dimensions through the CS (Cross-dimension Self-attention) mechanism. Meantime, every dimension is complementary to each other. Also, CS mechanism has the advantage of convolutions capturing local contextual weights. Finally, FCSU-Net is validated on several datasets, and the results show that FCSU-Net not only has a relatively small number of parameters, but also has a leading segmentation performance.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1879-0534  
入藏号:  
MEDLINE:39094324  
PubMed ID:  
39094324  
Indexed Date:  
2024-08-04  
Institution Address:  
School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China. Electronic address: xsj0410@mails.swust.edu.cn.  
Institution Address:  
School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China. Electronic address: c\_y\_f\_1@163.com.  
Institution Address:  
School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China. Electronic address: yangsk3230@163.com.  
Institution Address:  
School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China. Electronic address: zhxq0528@163.com.  
Institution Address:  
Mianyang Central Hospital, School of Medicine, University of Electronic Science and Technology of China, Mianyang 621000, China; NHC Key Laboratory of Nuclear Technology Medical Transformation, Mianyang Central Hospital, Mianyang 621000, China. Electronic address: sf266@qq.com.  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 357. Depth estimation of supervised monocular images based on semantic

摘要: In recent years, the research method of depth estimation of target images using Convolutional Neural Networks (CNN) has been widely recognized in the fields of artificial intelligence, scene understanding and three-dimensional (3D) reconstruction. The fusion of semantic segmentation information and depth estimation will further improve the quality of acquired depth images. However, how to deeply combine image semantic in-formation with image depth information and use image edge information more accurately to improve the ac-curacy of depth image is still an urgent problem to be solved. For this purpose, we propose a novel depth estimation model based on semantic segmentation to estimate the depth of monocular images in this paper. Firstly, a shared parameter model of semantic segmentation information and depth estimation information is built, and the semantic segmentation information is used to guide depth acquisition in an auxiliary way. Then, through the multi-scale feature fusion module, the feature information contained in the neural network on different layers is fused, and the local feature information and global feature information are effectively used to generate high-resolution feature maps, so as to achieve the goal of improving the quality of depth image by optimizing the semantic segmentation model. The experimental results show that the model can fully extract and combine the image feature information, which improves the quality of monocular depth vision estimation. Compared with other advanced models, our model has certain advantages.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
11  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
1047-3203  
eISSN:  
1095-9076  
入藏号:  
WOS:000993683400001  
Indexed Date:  
2023-06-01  
Institution Address:  
Changchun Univ Sci & Technol, Coll Elect & Informat Engn, Changchun 130022, Peoples R China  
Institution:

### 358. Lightweight concrete crack detection based on spiking neural networks

摘要: Purpose - Most existing methods for concrete crack detection are based on deep learning techniques such as convolutional neural networks. However, these models, due to their large memory footprint, high power consumption and insufficient feature extraction capabilities, face challenges in mobile applications. To address these issues, this paper proposes a lightweight spiking neural network detection model. Design/methodology/approach - This model achieves fast and accurate crack detection. Firstly, the Gabor-Spiking (GS) module preprocesses input images, extracting texture features and edge features of crack images through Gabor filter convolution modules and spiking convolution modules, respectively. Next, the multiscale residual (MR) module is designed, composed of convolutional layers and residual modules of various scales, to process the fused features and perform crack detection. Findings - Experimental results demonstrate that the model's size can be reduced to 4.6 MB, achieving accuracy improvements to 87.3 and 96.4% on the SDNET and OCD datasets, respectively. Originality/value - This paper proposes a lightweight spiking neural network detection model based on the GS module for edge texture feature fusion and the MR module for crack detection.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0264-4401  
eISSN:  
1758-7077  
入藏号:  
WOS:001345786900001  
Indexed Date:  
2024-11-11  
Institution Address:  
Guangdong Univ Technol, Sch Integrated Circuits, Guangzhou, Peoples R China  
Institution:

### 359. VcaNet: Vision Transformer with fusion channel and spatial attention

摘要: Accurate segmentation of brain tumors from MRI scans is a critical task in medical image analysis, yet it remains challenging due to the complex and variable nature of tumor shapes and sizes. Traditional convolutional neural networks (CNNs), while effective for local feature extraction, struggle to capture long-range dependencies crucial for 3D medical image analysis. To address these limitations, this paper presents VcaNet, a novel architecture that integrates a Vision Transformer (ViT) with a fusion channel and spatial attention module (CBAM), aimed at enhancing 3D brain tumor segmentation. The encoder of VcaNet employs a 3D enhanced convolution (ENCO) module to capture local volumetric features, while a Vision Transformer and multi-scale feature fusion module are incorporated in the bottleneck to capture global dependencies. Additionally, a CBAM is introduced in the decoder to further improve the integration of local and global features, enhancing segmentation accuracy. Extensive experiments on the two public BraTS Datasets demonstrate that VcaNet outperforms existing models, particularly in handling the complex spatial structures of brain tumors. This approach provides valuable insights for improving brain tumor segmentation, and its performance in 3D tasks surpasses that of 2D models, laying a foundation for future advancements in medical imaging.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
2  
eISSN:  
1879-0534  
入藏号:  
MEDLINE:39813745  
PubMed ID:  
39813745  
Indexed Date:  
2025-01-19  
Institution Address:  
College of Physics and Electronic Information Engineering, Zhejiang Normal University, Jinhua, 321004, China. Electronic address: pandc561@zjnu.edu.cn.  
Institution Address:  
College of Physics and Electronic Information Engineering, Zhejiang Normal University, Jinhua, 321004, China; Zhejiang Institute of Optoelectronics, Jinhua, 321004, China. Electronic address: shenjianguo@zjnu.cn.  
Institution Address:  
Stirling College, Chengdu University, Chengdu, Sichuan, 610106, China. Electronic address: zaid@stir.ed.cn.  
Institution Address:  
College of Physics and Electronic Information Engineering, Zhejiang Normal University, Jinhua, 321004, China; Zhejiang Institute of Optoelectronics, Jinhua, 321004, China. Electronic address: alqaness@zjnu.edu.cn.  
Institution:  
Institution:  
Institution:  
Institution:

### 360. Fault Diagnosis of Electric Submersible Pumps Using a Three-Stage

摘要: A convolutional neural network support vector machine (CNN-SVM) method based on multichannel feature fusion is used for progressive fault diagnosis of offshore oil and gas wells. The excellent classification performance of CNN is attributed to its ability to extract feature representations from large amounts of easily distinguishable data. However, the capability of CNN is severely constrained by the noisy and small sample amount of the electric submersible pump fault data to be studied in this article. First, 12 representative statistical features are extracted from the raw data to reduce the noise. Then, the feature mapping model is designed based on CNN migration learning. Finally, SVM is used instead of softmax function to adopt representative features directly from the mapping model for fault classification. Comparative experimental results show that the accuracy of fault diagnosis using feature-extracted samples is better than using the raw samples directly. The proposed CNN-SVM approach has the best classification results compared to SVM, BPNN, CNN, BPNN-SVM, CNN-Attention, CNN-LSTM, and CNN-LSTM-Attention, which implies that manual feature extraction is still an indispensable tool in the fault diagnosis process.  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
2194-4288  
eISSN:  
2194-4296  
入藏号:  
WOS:001036979100001  
Indexed Date:  
2023-08-08  
Institution Address:  
Guangdong Ocean Univ, Dept informat & Comp Sci, Zhanjiang 524088, Guangdong, Peoples R China  
Institution Address:  
Guangdong Med Univ, Educ Technol & Informat Ctr, Zhanjiang 524023, Guangdong, Peoples R China  
Institution Address:  
CNOOC China Ltd, ZhanJiang Branch, Zhanjiang 524057, Guangdong, Peoples R China  
Institution Address:  
Southern Marine Sci & Engn Guangdong Lab Zhanjiang, Zhanjiang 524088, Guangdong, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 361. Research on ZYNQ neural network acceleration method for aluminum surface

摘要: Convolutional Neural Networks (CNN) are an important means of detection of microdefects on the aluminum surface, and the high complexity and computing power requirements of the CNN model lead to difficulties in deploying them on edge computing platforms as the detection accuracy continues to improve. We have studied a lightweight acceleration method for detecting microdefects on aluminum surfaces on the Zynq-7000 All Programmable SoC (ZYNQ) platform. A lightweight aluminum surface defect detection network (LADFastDet) and high-performance accelerators based on ZYNQ are designed to meet the requirements of precision and speed under limited resources. In the LADFastDet structure, a lightweight inverted residual block is designed by combining depthwise convolution, inverted residual block, and inverted bottleneck. A multiscale feature fusion structure is designed to effectively improve the detection accuracy of LADFastDet, especially small target defects. We design accelerators on ZYNQ through optimization methods such as loop optimization strategy, ping-pong buffering, and multichannel and multiple interfaces data reading and writing to reduce data access latency and thus improve the computing speed. The experimental results show that the LADFastDet model has a mAP of 97.51%, the inference time of the accelerators for a single image is 42.57 ms, and a power consumption of 2.15 W, which achieves a throughput of 24.9 GOPS and an energy efficiency of 11.58 GOPS/W.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1051-2004  
eISSN:  
1095-4333  
入藏号:  
WOS:001372071200001  
Indexed Date:  
2024-12-13  
Institution Address:  
Changsha Univ Sci & Technol, Sch Phys & Elect Sci, Changsha 410114, Hunan, Peoples R China  
Institution:

### 362. A novel end-to-end chromosome classification approach using deep neural

摘要: The automated karyotyping analysis serves as a crucial tool in diagnosing genetic disorders stemming from chromosome abnormalities. Chromosome classification, a pivotal step in this diagnostic process encounters challenges due to the non -rigid structure and unpredictable curved shapes of chromosomes. Extracting features with discriminative power for enhanced classification accuracy becomes difficult, impacting the practical results of karyotyping analyses in clinical applications. In this paper, we propose a novel end -to -end chromosome classification approach based on deep neural network with a triple attention mechanism (TAM). The approach unfolds in three stages. Firstly, deep CNNs and modified Multiscale Feature Fusion (MFF) blocks collaborate to generate informative multi -scale global representations. In the second stage, to uncover finer features that enhance chromosome differentiation, TAM is devised. It learns channel -wise, spatial -wise and point -wise attention feature maps fully extracting valuable information from multi -scale global features. Finally, a probability -weighted prediction scheme is incorporated to synthetically estimate classification branches at all scales, facilitating type identification. The evaluation results, based on a clinical G -band chromosome dataset comprising 52,808 single chromosome images, indicate that the proposed chromosome classification approach attains an impressive accuracy of 99.4%. In conclusion, our method exhibits strong robustness and high stability in effectively classifying multiple types of chromosomes.  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:001162207700001  
Indexed Date:  
2024-02-22  
Institution Address:  
Shanghai Jiao Tong Univ, Dept Automat, Shanghai, Peoples R China  
Institution:

### 363. Object Detection Method Based on Multi-scale Feature Fusion for Driving

摘要: Aiming at the problem of low detection accuracy of convolutional neural network model for object detection in driving vision, a multi-scale feature fusion object detection method based on improved RefineDet is proposed.Firstly, the LFIP(Light-weight Featured Image Pyramid)network is embedded in the RefineDet, and the multi-scale feature map generated by LFIP network is integrated with the main feature map output from ARM(Anchor Refinement Module)in the RefineDet, which improves the output effect of anchors preliminary classification and regression in the convolutional layer, and provides refined anchors frame for ODM(Object Detection Module)for further regression and multi-class prediction.Secondly, after the ODM in the RefineDet, a multi-branch structure RFB(Receptive Field Block)is embedded to obtain receptive fields of different scale in the detection task to improve the features extracted from the backbone network.Thirdly, the activation function in the model is replaced by the nonlinear activation function PReLU(Parametric Rectified Linear Unit)with learnable parameters to speed up the convergence of the model.Then, the Bounding box loss function of RefineDet is replaced by the Repulsion Loss function to narrow the gap between a proposal and its designated target and increase the distance between the proposal and the surrounding non-target objects.Finally, an object detection dataset is constructed with 48 260 images in driving vision, including 38 608 as training set and 9 652 as test set, which are verified on mainstream GPU hardware platform.The mAP of this method is 85.59%, which is better than RefineDet and other improved algorithms;the FPS is 41.7 frame/s, which meets the application requirements of driving scene object detection.Experimental results show that the proposed method can improve the accuracy of object detection in driving vision, and solve the problems of occlusion object detection and small object detectionin driving vision to a certain extent.  
摘要:  
针对驾驶场景中目标检测卷积神经网络模型检测精度较低的问题,提出一种基于改进RefineDet网络结构的多尺度特征融合目标检测方法。在RefineDet网络结构中嵌入LFIP(Light-weight Featurized Image Pyramid,轻量级特征化的图像金字塔)网络,将LFIP网络生成的多尺度特征图与RefineDet中的ARM(Anchor Refinement Module,锚点框修正模块)输出的主特征图相融合,提升特征层中锚点框初步分类和回归的输出效果,为ODM(Object Detection Module,目标检测模块)模块提供修正的锚点框以便于进一步回归和多类别预测;在RefineDet网络结构中的ODM之后嵌入多分支结构RFB(Receptive Field Block,感受野模块),在检测任务中获得不同尺度的感受野以改善主干网络中提取的特征。将模型中的激活函数替换为带有可学习参数的非线性激活函数PReLU(Parametric Rectified Linear Unit,参数化修正线性单元),加快网络模型的收敛速度;将RefineDet的边界框回归损失函数替换为排斥力损失函数Repulsion Loss,使目标检测中的某预测框更靠近其对应的目标框,并使该预测框远离附近的目标框及预测框,可以提升遮挡情况下目标检测的精度;构建驾驶视觉下的目标检测数据集,共计48 260张,其中38 608张作为训练集,9 652张作为测试集,并在主流的GPU硬件平台进行验证。该方法的mAP为85.59%,优于RefineDet及其他改进算法;FPS为41.7 frame/s,满足驾驶场景目标检测的应用要求。实验结果表明,该方法在检测速度略微下降的情况,能够较好地提升驾驶视觉下的目标检测的精确度,并能够一定程度上解决驾驶视觉下的遮挡目标检测和小目标检测的问题。  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1002-8331  
入藏号:  
CSCD:7029657  
Indexed Date:  
2021-10-30  
Institution Address:  
华南理工大学电子与信息学院, 广州, 广东 510640, 中国  
Institution Address:  
广东白云学院大数据与计算机学院, 广州, 广东 510450, 中国  
Institution Address:  
广州市生发科技服务有限公司技术部, 广州, 广东 510308, 中国  
Institution Address:  
School of Electronic and Information Engineering, South China University of Technology, Guangzhou, Guangdong 510640, China  
Institution Address:  
Faculty of Mega Data and Computer Science, Guangdong Baiyun University, Guangzhou, Guangdong 510450, China  
Institution Address:  
Department of Technology, Guangzhou Shengfa Technology Service Co., Ltd., Guangzhou, Guangdong 510308, China  
Institution:  
华南理工大学电子与信息学院  
Institution:  
广东白云学院大数据与计算机学院  
Institution:  
广州市生发科技服务有限公司技术部  
Institution:  
School of Electronic and Information Engineering, South China University of Technology  
Institution:  
Faculty of Mega Data and Computer Science, Guangdong Baiyun University  
Institution:  
Department of Technology, Guangzhou Shengfa Technology Service Co., Ltd.

### 364. Hyperspectral image classification based on deep separable residual

摘要: Hyperspectral image have rich spatial and spectral information, and how to fully extract and utilize the features of these two dimensions is a research hotspot in hyperspectral classification methods. At present, the unique convolutional operation and deep feature extraction structure of convolutional neural network enable them to have stronger feature representation capabilities and achieve good results in hyperspectral image classification. However, CNN methods do not assign different weights based on the importance of features in the feature extraction process, making it difficult to effectively utilize key features, and most importantly, using fixed shaped convolution kernel can easily overlook the differences between hyperspectral image features. A hyperspectral image classification method based on deep separable residual attention network is proposed to address the above issues. Firstly, to reduce the correlation between hyperspectral image data and minimize the interference of redundant information, principal component analysis is used to reduce the dimensionality of hyperspectral image. Secondly, a shallow feature extraction module is constructed, which can dynamically adjust the size of the receptive field according to the actual situation of the image, adaptively extract shallow features, and reduce the loss of original image features. Then, a depthwise separable residual attention mechanism module is proposed, based on which features are extracted. Starting from global and local features, contextual information on image features in channel and spatial domains is extracted. Finally, use a multi -scale feature fusion module to fully integrate feature maps at different scales. Using Indian Pines, Pavia University and Botswana as experimental datasets, the overall classification accuracy of this paper ' s method is 98.47 %, 98.70 %, 98.83 % with only 50, 50, 30 training samples per class. The Kappa coefficient is 98.25 %, 98.27 %, and 98.73 %, respectively. Compared with advanced methods, this method not only has higher classification accuracy, but also fully utilizes key features at various network levels.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1350-4495  
eISSN:  
1879-0275  
入藏号:  
WOS:001259161400001  
Indexed Date:  
2024-07-08  
Institution Address:  
Liaoning Tech Univ, Sch Geomat, Fuxin 123000, Peoples R China  
Institution Address:  
Liaoning Tech Univ, Sch Software, Huludao 125105, Peoples R China  
Institution:  
Institution:

### 365. Investigating transformer-based models for spatial downscaling and

摘要: High-resolution and accurate prediction of near-surface weather parameters based on numerical weather prediction (NWP) models is essential for many downstream and real-world applications. Traditional dynamical or statistical downscaling methods are insufficient to derive high-resolution data from operational NWP forecasts, making it essential to devise new approaches. In recent years, an increasing number of researchers have explored the implementations of deep learning (DL) based models for spatial downscaling, motivated by the similarity between the super-resolution (SR) problem in computer vision (CV) and downscaling. Furthermore, while transformer-based models have become state-of-the-art models for many SR tasks, they are rarely applied for downscaling of weather forecasts or climate projections. This study adapted transformer-based models such as SwinIR and Uformer to downscale the temperature at 2 m (T2m$$ {T}\_{2\mathrm{m}} $$) and wind speed at 10 m (WS10m$$ W{S}\_{10\mathrm{m}} $$) over Eastern Inner Mongolia, encompassing the area from 39.6-46 degrees N latitude and 111.6-118 degrees E longitude. We used high-resolution forecast (HRES) data from the European Centre for Medium-range Weather Forecast (ECMWF) with a spatial resolution of 0.1 degrees as the input and gridded observation data from the China Meteorological Administration (CMA) Land Data Assimilation System (CLDAS) at a spatial resolution of 0.01 degrees as the target. Given that the models use observation data rather than a coarse-grained version of forecast data as the target, they accomplish both bias correction and spatial downscaling. The results demonstrate that the performance of SwinIR and Uformer is superior to that of two convolutional neural network (CNN) based models (UNet and RCAN). Additionally, we introduced a novel module to extract features of varying resolution from the high-resolution topography data and applied a multiscale feature fusion module to merge features of different scales, contributing to further enhancement of Uformer's performance.  
For the first time, transformer-based models such as SwinIR and Uformer are applied for weather downscaling. Comparison with CNN-based models demonstrates that transformer-based models outperform CNN-based models such as UNet and RCAN. Furthermore, topography feature extraction and multiscale feature fusion modules are proposed to improve Uformer's downscaling performance.image  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
0035-9009  
eISSN:  
1477-870X  
入藏号:  
WOS:001089034600001  
Indexed Date:  
2023-11-08  
Institution Address:  
Damo Acad, Alibaba Grp, Hangzhou, Peoples R China  
Institution Address:  
Fudan Univ, Artificial Intelligence Innovat & Incubat Inst, Shanghai, Peoples R China  
Institution Address:  
Fudan Univ, Artificial Intelligence Innovat & Incubat Inst, Shanghai 200433, Peoples R China  
Institution:  
Institution:  
Institution:

### 366. Strip Steel Surface Quality Detection System Based on YOLOv5

摘要: For identifying and classifying surface defects of strip steel in industrial production and improving the efficiency of defect detection, a detection and recognition system for surface quality defects of strip steel in industrial production environments is designed in this article. Firstly, industrial raw material strip steel was used as the research object, and an experimental operating environment was established to label the dataset. Then, entered into the image, image scaled by YOLOv5s, and then the image is scaled into S x S using convolutional neural networks. Feature maps are extracted from each small grid. Secondly, the extracted feature maps were pooled and subjected to a series of sharding, stitching, and multi-scale feature fusion. Finally, under the GPU acceleration of the PyTorch learning framework, the processing speed was improved by identifying categories through a fully connected layer network, all indicators have been detected, as well as information on the category and confidence level of defects identified. The experimental results show that the system can identify four types of defects on the surface of the strip steel and achieve an accuracy of 90% in identifying defect types.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1058-4587  
eISSN:  
1607-8489  
入藏号:  
WOS:001344537600008  
Indexed Date:  
2024-11-09  
Institution Address:  
Xian Univ, Shaanxi key Lab Surface Engn & Remfg, Xian 710065, Shaanxi, Peoples R China  
Institution Address:  
Xian Univ, Xian Key Lab Prototyping & Optimizat Implantable D, Xian, Peoples R China  
Institution:  
Institution:

### 367. MSF-Model: Multi-Scale Feature Fusion-Based Domain Adaptive Model for

摘要: One of the most common causes of mortality for women globally is breast cancer. Early breast cancer identification could make it possible for people to receive the appropriate treatment to save their lives and return to their routine lives. Breast cancer diagnosis by histopathology is referred to as the gold standard. In recent years, convolutional neural network-based techniques are used for breast cancer classification. However, they faced domain adaptation, small objects retention, and feature extraction issues of complex microscopic images. In this study, we introduced multi-scale feature fusion-based domain adaptive model for breast cancer classification using histopathology images. It has two blocks and six lightweight sub-models where each block contains three models. Dilated layers are used in sub-models to overcome the disappearing of small objects in deep layers. Reducing the disappearing of small objects helped to extract better features for higher performance. Multiple heterogeneous feature extractors are used in this study which helped to extract various features. Extracted features are fused and reduced by retaining better features. Learning of model from natural images to complex microscopic images has limitation of domain adaptation. Same domain transfer learning is used in this study to overcome the limitations of different domain transfer learning. Model is trained on patchcamelyon17 dataset and weights of this training are further used for same domain transfer learning. Pre-trained weights are further used for the training of proposed model on BreaKHis dataset. A number of conventional data augmentation techniques are used as complex models require higher number of samples for the tuning of weights. Local window based CLAHE contrast enhancement technique is used to increase foreground-background contrast and remove noise. The proposed model achieved 98.00% precision, 98.15% recall, 98.08% f-measure, and 98.23% accuracy on test data. To best of our knowledge, it surpassed state-of-the-art models.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
2169-3536  
入藏号:  
WOS:000892887200001  
Indexed Date:  
2022-12-14  
Institution Address:  
COMSATS Univ Islamabad CUI, Dept Comp Sci, Islamabad 45550, Pakistan  
Institution Address:  
Women Univ Swabi, Dept Comp Sci, Swabi 23430, Pakistan  
Institution Address:  
Seoul Natl Univ, Sch Elect & Comp Engn, Seoul 08826, South Korea  
Institution Address:  
King Khalid Univ, Dept & Coll Comp Sci, Abha 62529, Saudi Arabia  
Institution:  
Institution:  
Women Univ Swabi  
Institution:  
Institution:

### 368. YOGA: Deep object detection in the wild with lightweight feature

摘要: We introduce YOGA, a deep learning based yet lightweight object detection model that can operate on low-end edge devices while still achieving competitive accuracy. The YOGA architecture consists of a two-phase feature learning pipeline with a cheap linear transformation, which learns feature maps using only half of the convolution filters required by conventional convolutional neural networks. In addition, it per -forms multi-scale feature fusion in its neck using an attention mechanism instead of the naive concate-nation used by conventional detectors. YOGA is a flexible model that can be easily scaled up or down by several orders of magnitude to fit a broad range of hardware constraints. We evaluate YOGA on COCO-val and COCO-testdev datasets with over 10 state-of-the-art object detectors. The results show that YOGA strikes the best trade-off between model size and accuracy (up to 22% increase of AP and 23-34% reduc-tion of parameters and FLOPs), making it an ideal choice for deployment in the wild on low-end edge devices. This is further affirmed by our hardware implementation and evaluation on NVIDIA Jetson Nano.(c) 2023 Elsevier Ltd. All rights reserved.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
被引频次合计:  
7  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:000953091000001  
Indexed Date:  
2023-04-05  
Institution Address:  
Missouri Univ Sci & Technol, Dept Comp Sci, Rolla, MO 65409 USA  
Institution:

### 369. A finger vein authentication method based on the lightweight Siamese

摘要: Finger vein authentication is an efficient and convenient biometric verification technology which has been widely investigated in personal identification. However, the vein image quality can be reduced by changeable finger postures and near-infrared light distribution during the finger vein collection procedure. The poor finger vein images tend to cause feature loss, affecting final authentication results. Therefore, we propose a lightweight Siamese network with a self-attention mechanism to improve the authentication performance for low-quality finger vein images. Firstly, the finger vein features are extracted from a three-layer convolutional neural network. Then, we introduce the global context network to model the feature maps from different layers, obtaining global information about the vein features. Next, a multi-scale feature fusion strategy is proposed to fuse the feature maps from different scales, which increases the diversity of vein features. Finally, the self-attention convolution module is proposed to weight and vectorize the fused features. Extensive experiments are conducted on three datasets, MMCBNU\_6000, SDUMLA-HMT, and FV\_USM, demonstrating that the proposed method can significantly improve over the state-of-the-art methods.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
9  
ISSN:  
1350-4495  
eISSN:  
1879-0275  
入藏号:  
WOS:000911028400001  
Indexed Date:  
2023-01-27  
Institution Address:  
Heilongjiang Univ, Coll Elect Engn, Harbin 150080, Peoples R China  
Institution:

### 370. Semi-finished flywheel disk based on deep learning research on surface

摘要: The surface defects of flywheel disc semi-finished products have complex and changeable morphological characteristics and random distribution. At present, relevant enterprises can only detect them through manual visual inspection. However, the low efficiency of manual inspection and the unstable inspection quality can easily lead to false inspections and missed inspections, which cannot meet the growing demand for production capacity. In order to achieve intelligent and efficient detection of defects, this paper proposes a surface defect detection algorithm for flywheel disc semi-finished products based on improved faster region-based convolutional neural networks (Faster R-CNN). First of all, based on multi-scale feature fusion, residual feature recalibration and deformable convolution, this paper designs a feature extraction network that can better capture and characterize defect morphology. Secondly, optimize the design of Faster R-CNN algorithm, use k-means++ cluster analysis to optimize the anchor generation rules in the network, so as to adapt to the defects of large aspect ratio, the region of interest (ROI) pooling calculation method incorporating global feature information is redesigned to prevent the position deviation of candidate areas when they are mapped back to the original image. Aiming at the problem that adjacent overlapping positive samples are deleted by mistake, the soft non-maximum suppression (Soft-NMS) algorithm is used to optimize the non-maximum suppression process and increase the number of positive samples output by the region proposal network(RPN). Then, the surface defect images are collected to build a data set, aiming at the problem that the data set is small and the distribution of the number of defects in each category is unbalanced, the classical data enhancement methods are used to augment the data set and equalize the defect categories. Finally, the surface defect detection and application experiment research of flywheel disc semi-finished products is carried out. The detection accuracy of the algorithm in this paper on the surface defect test set reaches 92.7%, which is 9.6% higher than the original Faster R-CNN detection accuracy, and 18.5% higher for the detection accuracy of small minor defects, and the improvement effect is more obvious.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
2051-672X  
入藏号:  
WOS:000862013500001  
Indexed Date:  
2022-10-09  
Institution Address:  
Changchun Univ Sci & Technol, Joint Recearch Ctr Diamond Tool Technol, Changchun 130022, Peoples R China  
Institution Address:  
Changchun Univ Sci & Technol, Chongqing Res Inst, Chongqing 401135, Peoples R China  
Institution:  
Institution:

### 371. Hybrid SFNet Model for Bone Fracture Detection and Classification Using

摘要: An expert performs bone fracture diagnosis using an X-ray image manually, which is a time-consuming process. The development of machine learning (ML), as well as deep learning (DL), has set a new path in medical image diagnosis. In this study, we proposed a novel multi-scale feature fusion of a convolution neural network (CNN) and an improved canny edge algorithm that segregate fracture and healthy bone image. The hybrid scale fracture network (SFNet) is a novel two-scale sequential DL model. This model is highly efficient for bone fracture diagnosis and takes less computation time compared to other state-of-the-art deep CNN models. The innovation behind this research is that it works with an improved canny edge algorithm to obtain edges in the images that localize the fracture region. After that, grey images and their corresponding canny edge images are fed to the proposed hybrid SFNet for training and evaluation. Furthermore, the performance is also compared with the state-of-the-art deep CNN models on a bone image dataset. Our results showed that SFNet with canny (SFNet + canny) achieved the highest accuracy, F1-score and recall of 99.12%, 99% and 100%, respectively, for bone fracture diagnosis. It showed that using a canny edge algorithm improves the performance of CNN.  
Web of Science 核心合集中的 "被引频次":  
26  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
26  
eISSN:  
1424-8220  
入藏号:  
WOS:000839767000001  
PubMed ID:  
35957380  
Indexed Date:  
2022-08-22  
Institution Address:  
GLA Univ, Dept Comp Engn & Applicat, Mathura 281406, Uttar Pradesh, India  
Institution Address:  
Koneru Lakshmaiah Educ Fdn, Dept Comp Sci & Engn, Vaddeswaram 522502, Andhra Pradesh, India  
Institution Address:  
Chitkara Univ, Sch Engn & Technol, Baddi 174103, Himachal Prades, India  
Institution Address:  
Dar Al Hekma Univ, Hekma Sch Engn Comp & Informat, Dept Comp Sci, Jeddah 222464872, Saudi Arabia  
Institution Address:  
Univ Turku, Dept Comp, Turku 20500, Finland  
Institution Address:  
Univ Monastir, Higher Inst Comp Sci & Math, Dept Technol, Monastir 5000, Tunisia  
Institution:  
Institution:  
Institution:  
Chitkara Univ  
Institution:  
Institution:  
Institution:

### 372. MHIAIFormer: Multihead Interacted and Adaptive Integrated Transformer

摘要: Deep learning is an effective method for hyperspectral image (HSI) classification, where CNN-based and Transformer-based methods have achieved excellent performance. However, there are some drawbacks to the existing CNN-based and Transformer-based HSI classification approaches: 1) CNN-based methods are deficient in showing the extraction of multiscale features and localized features owing to the fixed-size input patch. 2) the MHSA module ignores the interaction capability between multiple attention heads, which leads to insufficient feature fusion in various directions. 3) The weights of attention heads in various directions are disregarded in the MHSA and attention heads are simply concatenated horizontally. To address the above-mentioned limitations, a novel multihead interacted and adaptive integrated transformer (MHIAIFormer) with spatial-spectral attention, which integrates the respective advantages of convolutions and transformers is proposed in this study. A pyramidal spatial-spectral attention (PS2A) feature extraction module is adopted to efficiently capture the localized and multiscale feature information of HSI. The output of PS2A is then sent to the transformer encoder stage through a grouped multiscale cross-dimension embedding module, which includes additive self-attention using multihead interaction and MHSA with adaptive multihead merging to capture the long-range dependencies of the features. Extensive experiments on four datasets verify that our proposed approach achieves more satisfactory classification accuracy when compared with state-of-the-art models. The overall accuracy of the proposed model achieved 95.97%, 98.68%, 92.68%, and 99.49% on four datasets.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001303437700001  
Indexed Date:  
2024-09-16  
Institution Address:  
Qingdao Univ, Coll Comp Sci & Technol, Remote Sensing Informat & Digital Earth Ctr, Qingdao 266071, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Bangabandhu Sheikh Mujibur Rahman Agr Univ, Dept Agr Extens & Rural Dev, Gazipur 1706, Bangladesh  
Institution:  
Institution:  
Institution:

### 373. HT-Net: A Hybrid Transformer Network for Fundus Vessel Segmentation

摘要: Doctors usually diagnose a disease by evaluating the pattern of abnormal blood vessels in the fundus. At present, the segmentation of fundus blood vessels based on deep learning has achieved great success, but it still faces the problems of low accuracy and capillary rupture. A good vessel segmentation method can guide the early diagnosis of eye diseases, so we propose a novel hybrid Transformer network (HT-Net) for fundus imaging analysis. HT-Net can improve the vessel segmentation quality by capturing detailed local information and implementing long-range information interactions, and it mainly consists of the following blocks. The feature fusion block (FFB) is embedded in the shallow levels, and FFB enriches the feature space. In addition, the feature refinement block (FRB) is added to the shallow position of the network, which solves the problem of vessel scale change by fusing multi-scale feature information to improve the accuracy of segmentation. Finally, HT-Net's bottom-level position can capture remote dependencies by combining the Transformer and CNN. We prove the performance of HT-Net on the DRIVE, CHASE\_DB1, and STARE datasets. The experiment shows that FFB and FRB can effectively improve the quality of microvessel segmentation by extracting multi-scale information. Embedding efficient self-attention mechanisms in the network can effectively improve the vessel segmentation accuracy. The HT-Net exceeds most existing methods, indicating that it can perform the task of vessel segmentation competently.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
7  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
eISSN:  
1424-8220  
入藏号:  
WOS:000859551700001  
PubMed ID:  
36146132  
Indexed Date:  
2022-10-01  
Institution Address:  
Xinjiang Univ, Coll Informat Sci & Engn, Urumqi 830000, Peoples R China  
Institution:

### 374. A multi-scale temporal convolutional capsule network with parameter-free

摘要: We proposed a multi-scale temporal convolutional capsule network model coupled with a parameter-free attention module and dynamic routing mechanism to analyze complex vibration signals for diagnosing the health of bearings. The proposed method utilizes a capsule network as the fundamental architecture. Instead of a convolutional neural network, a temporal convolutional network is employed. Additionally, a multi-scale feature fusion module is integrated into the capsule network structure to dynamically extract various layers of features from fault samples, enhancing the discriminatory capability of abnormal data. Subsequently, the parameter-free attention module and dynamic routing mechanism are employed to construct digital capsules. This allows the smallest unit capsule in a single layer to carry more information, enhance the similarity between the instance primary capsule and the fault capsule, reduce the interference of irrelevant features to the model, and improve the accuracy of fault type recognition. Finally, a multi-scale temporal convolutional capsule network model that integrates feature extraction and pattern recognition is established to perform end-to-end diagnosis of the bearing. Experimental findings suggest that the proposed method outperforms other deep learning methods in terms of accuracy and robustness. It can provide a theoretical basis and implementation path for the detection and diagnosis of train wheelset bearing time series abnormal data.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001360103300001  
Indexed Date:  
2024-11-28  
Institution Address:  
Southwest Jiaotong Univ, Sch Mech Engn, Chengdu 610031, Peoples R China  
Institution Address:  
Univ Elect Sci & Technol China, Sch Automat Engn, Chengdu 611731, Peoples R China  
Institution Address:  
Harbin Inst Technol, Sch Astronaut, Harbin 150001, Peoples R China  
Institution:  
Institution:  
Institution:

### 375. ADNet: A Real-Time Floating Algae Segmentation Using Distillation

摘要: The segmentation of floating algae is a hot topic in the field of marine environmental research. Given the vastness of coastal areas and complex environments, algae detection models must have both higher performance and lower deployment costs. However, relying solely on a single Convolutional Neural Network (CNN) or transformer structure fails to achieve this objective. In this paper, a novel real-time floating algae segmentation method using a distillation network (ADNet) is proposed, based on the RGB images. ADNet can effectively transfer the performance of the transformer-based teacher network to the CNN-based student model while preserving its lightweight design. Faced with complex marine environments, we introduce a novel Channel Purification Module (CPM) to simultaneously strengthen algae features and purify interference responses. Importantly, the CPM achieves this operation without increasing any learnable parameters. Moreover, considering the huge scale differences among algae targets in surveillance RGB images, we propose a lightweight multi-scale feature fusion network (L-MsFFN) to improve the student's modeling ability across various scales. Additionally, to mitigate interference from low-level noises on higher-level semantics, a novel position purification module (PPM) is proposed. The PPM can achieve more accurate weight attention calculation between different pyramid levels, thereby enhancing the effectiveness of fusion. Compared to CNNs and transformers, our ADNet strikes an optimal balance between performance and speed. Extensive experimental results demonstrate that our ADNet achieves higher application performance in the field of floating algae monitoring tasks.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2077-1312  
入藏号:  
WOS:001256509700001  
Indexed Date:  
2024-07-04  
Institution Address:  
MNR, East China Sea Ecol Ctr, Shanghai 201206, Peoples R China  
Institution Address:  
Shanghai East Sea Marine Engn Survey & Design Inst, Shanghai 200136, Peoples R China  
Institution Address:  
State Ocean Adm, East Sea Informat Ctr, Shanghai 200136, Peoples R China  
Institution:  
MNR  
Institution:  
Shanghai East Sea Marine Engn Survey & Design Inst  
Institution:  
State Ocean Adm

### 376. Traffic Sign Detection and Recognition Using Multi-Scale Fusion and

摘要: Traffic sign detection, though one of the key technologies in intelligent transportation, still has bottleneck in accuracy due to the small size and diversity of traffic signs. To solve this problem, we proposed a two-stage CNN object detection algorithm based on multi-scale feature fusion and prime sample attention. We improved the original Faster R-cnn model in terms of feature extraction and sampling strategy. For feature extraction, to elevate the ability of neural networks to detect small objects, we adopted HRNet as the feature extractor. There are four stages in HRNet - a series of high resolution subnets as the starting point with repeated adding parallel high to low resolution subnets to form other stages. In the whole process, the information in the parallel multi-resolution sub-network is repeatedly exchanged to perform repeated multi-scale fusion. For sampling strategy, we adopted a simple and effective sampling and learning strategy called Prime Sample Attention (PISA), consisting of Importance-based Sample Reweighting (ISR) and Classification Aware Regression Loss (CARL). PISA proposed the concepts of IoU Hierarchical Partial Sorting (IoU-HLR) and Hierarchical Partial Score Sorting (Score-HLR), which sort the importance of positive samples and negative samples in mini-batch respectively. With the proposed method, the training process is focusing on prime samples rather than evenly treat all ones. The algorithm complexity of our method is lower than that of other state-of-the-art. After experiments by TT100K dataset, our method can attain a comparable or even better detection accuracy and robustness.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
19  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
19  
ISSN:  
2169-3536  
入藏号:  
WOS:000606560700001  
Indexed Date:  
2021-07-21  
Institution Address:  
Nanjing Univ Sci & Technol, Sch Elect & Opt Engn, Nanjing 210094, Peoples R China  
Institution:

### 377. Small Sample Hyperspectral Image Classification Method Based on

摘要: Deep learning has achieved significant success in the field of hyperspectral image (HSI) classification, but challenges are still faced when the number of training samples is small. Feature fusing approaches based on multi-channel and multi-scale feature extractions are attractive for HSI classification where few samples are available. In this paper, based on feature fusion, we proposed a simple yet effective CNN-based Dual-channel Spectral Enhancement Network (DSEN) to fully exploit the features of the small labeled HSI samples for HSI classification. We worked with the observation that, in many HSI classification models, most of the incorrectly classified pixels of HSI are at the border of different classes, which is caused by feature obfuscation. Hence, in DSEN, we specially designed a spectral feature extraction channel to enhance the spectral feature representation of the specific pixel. Moreover, a spatial-spectral channel was designed using small convolution kernels to extract the spatial-spectral features of HSI. By adjusting the fusion proportion of the features extracted from the two channels, the expression of spectral features was enhanced in terms of the fused features for better HSI classification. The experimental results demonstrated that the overall accuracy (OA) of HSI classification using the proposed DSEN reached 69.47%, 80.54%, and 93.24% when only five training samples for each class were selected from the Indian Pines (IP), University of Pavia (UP), and Salinas Scene (SA) datasets, respectively. The performance improved when the number of training samples increased. Compared with several related methods, DSEN demonstrated superior performance in HSI classification.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
eISSN:  
2079-9292  
入藏号:  
WOS:000847031100001  
Indexed Date:  
2022-09-06  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Comp Sci, Natl Pilot Software Engn Sch, Beijing 100876, Peoples R China  
Institution:

### 378. MSFT-YOLO: Improved YOLOv5 Based on Transformer for Detecting Defects of

摘要: With the development of artificial intelligence technology and the popularity of intelligent production projects, intelligent inspection systems have gradually become a hot topic in the industrial field. As a fundamental problem in the field of computer vision, how to achieve object detection in the industry while taking into account the accuracy and real-time detection is an important challenge in the development of intelligent detection systems. The detection of defects on steel surfaces is an important application of object detection in the industry. Correct and fast detection of surface defects can greatly improve productivity and product quality. To this end, this paper introduces the MSFT-YOLO model, which is improved based on the one-stage detector. The MSFT-YOLO model is proposed for the industrial scenario in which the image background interference is great, the defect category is easily confused, the defect scale changes a great deal, and the detection results of small defects are poor. By adding the TRANS module, which is designed based on Transformer, to the backbone and detection headers, the features can be combined with global information. The fusion of features at different scales by combining multi-scale feature fusion structures enhances the dynamic adjustment of the detector to objects at different scales. To further improve the performance of MSFT-YOLO, we also introduce plenty of effective strategies, such as data augmentation and multi-step training methods. The test results on the NEU-DET dataset show that MSPF-YOLO can achieve real-time detection, and the average detection accuracy of MSFT-YOLO is 75.2, improving about 7% compared to the baseline model (YOLOv5) and 18% compared to Faster R-CNN, which is advantageous and inspiring.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
7  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
5  
Web of Science 核心合集中的 "被引频次":  
158  
被引频次合计:  
166  
eISSN:  
1424-8220  
入藏号:  
WOS:000796182900001  
PubMed ID:  
35591155  
Indexed Date:  
2022-05-26  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Modern Post, Beijing 100876, Peoples R China  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Artificial Intelligence, Beijing 100876, Peoples R China  
Institution Address:  
Beijing Univ Posts & Telecommun, Teaching Affairs Off, Beijing 100876, Peoples R China  
Institution:  
Institution:  
Institution:

### 379. Multi-scale representation for image deraining with state space model

摘要: With the rapid development of deep learning, significant progress has been made in image deraining techniques, with current methods primarily based on Convolutional Neural Network (CNN) and Transformer architectures. However, these approaches often face limitations due to the fixed receptive fields of CNN and the computational complexity of Transformer. Moreover, most existing image deraining methods adopt a single-input single-output network architecture, which struggles with multi-scale feature representation, particularly in capturing both global and local information, leading to potential information loss. To address these issues, this paper proposes a multi-scale visual state space model for image deraining, aiming to improve deraining performance and image restoration quality by combining multi-scale networks with state space model. Specifically, we design a multi-scale Mamba block that models global features with linear computational complexity, and we develop an efficient multi-scale 2D scanning strategy that uses geometric transformations to apply different numbers of scanning directions at various scales, thereby better extracting feature information at each scale. Additionally, we introduce a Frequency Feature Enhancement Module to capture local feature information, and a Gated Feature Fusion Module to adaptively aggregate complementary features across scales, further enhancing image restoration quality and deraining performance. Experimental results demonstrate that our method achieves superior deraining performance on multiple public benchmark datasets, outperforming the current state-of-the-art methods, while significantly improving efficiency and maintaining low computational cost.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001389241600006  
Indexed Date:  
2025-01-09  
Institution Address:  
Shenyang Aerosp Univ, Coll Elect & Informat Engn, Shenyang 110136, Peoples R China  
Institution Address:  
Dalian Maritime Univ, Coll Informat & Sci Technol, Dalian 116026, Peoples R China  
Institution:  
Institution:

### 380. CrackUnet: a novel network with joint network-in-network structure and

摘要: Automatic pavement crack detection is a critical technique in the intelligent transportation system, which can effectively replace the person with a machine to detect the pavement crack automatically. This task is excellently challenging due to the tiny texture and the various shapes of each crack object. Previous crack detection networks mainly aim to perform complicated multi-scale feature fusion to learn the semantic information of cracks. However, the typical symmetrical networks with high-to-low resolution are undesirable to extract detailed crack texture information, and these existing methods cannot effectively deal with the issue of the various shapes.This paper proposes a novel end-to-end U-shaped convolutional neural network, termed CrackUNet, for the pavement crack detection task. To extract the information of the tiny texture, we design a novel network-in-network structure, which can enlarge the receptive field and obtain multi-scale features by putting a sub-network into each convolutional layer. Besides, to handle the issue of the various shapes for the crack object, we exploit the deformable convolution to capture contextual information of each crack, which can further improve the performance of crack detection. We train and evaluate the proposed CrackUNet on three public pavement crack datasets. The quantitative experimental results illustrate that our network outperforms the current state-of-the-art methods with almost the same efficiency. Specifically, the precision, recall, and F1-score of the CrackUNet are approximately 92.27%, 93.99%, and 92.94%, respectively.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1868-8071  
eISSN:  
1868-808X  
入藏号:  
WOS:001132456100001  
Indexed Date:  
2024-01-10  
Institution Address:  
Southeast Univ, Sch Comp Sci & Engn, Nanjing 210019, Jiangsu, Peoples R China  
Institution:

### 381. Detection of railway object intrusion under infrared low light based on

摘要: There are some problems in railway object intrusion detection in infrared weak light environment, such as insufficient target feature extraction, low detection accuracy and real-time performance. Aiming to those problems, an anchor-free object intrusion depth learning model based on CenterNet target detection model is proposed. This model work with multi-feature fusion and attention enhancement. Firstly, based on the multi-scale feature extraction of infrared targets, the adaptive spatial feature fusion (ASFF) module is used for feature extraction. And to improve the feature extraction ability of infrared targets, this model makes full use of target high-level semantics and low-level fine-grained feature information. Secondly, the key features are extracted through the proposed modified dilated-convolutional block attention module (Dilated-CBAM), which expands the receptive field range of the attention mechanism module. On the one hand, this improvement overcomes the problem that the mapping area of the convolution block receiving field of the original central network becomes narrow and cannot detect weak and small targets; on the other hand, this improvement improves the detection accuracy of the anchor free network. Then, Smooth L1 loss function is used for training, which overcomes the problems of slow convergence speed and unstable solution of L1 loss function in the network training process. Finally, the experimental results are obtained through railway infrared data set and field experiments. The experimental results show that compared with the original CenterNet model, the average detection accuracy of this method is improved by 8.03%, the confidence of the detection frame is improved by 31.23%, and the average detection rate is 9.6 times higher than that of the Faster RCNN model. This method can detect railway object intrusion more quickly and accurately in the infrared weak light environment, both subjective and objective evaluation are better than the comparison method.  
摘要:  
针对红外弱光环境下铁路异物检测时存在目标特征提取不充分、检测精度及实时性低的问题,在CenterNet目标检测模型的基础上,提出了一种红外弱光下多特征融合与注意力增强的无锚框异物检测深度学习模型。在红外目标多尺度特征提取的基础上,引入自适应特征融合(ASFF)模块,充分利用目标高层语义与底层细粒度特征信息,提升红外目标特征提取能力。通过提出的空洞卷积增强注意力模块(Dilated-CBAM)进行关键特征提取,扩大注意力模块感受野范围,克服了原始CenterNet卷积块感受野映射区域变窄、无法检测弱小目标的问题,提升了无锚框网络的检测精度。使用Smooth L1损失函数进行训练,克服了L1损失函数在网络训练过程收敛速度慢及训练不稳定解的问题。通过铁路红外数据集及现场实验测试,结果表明:所提方法较原始CenterNet模型平均检测精度提高了8.03%,检测框置信度提升了31.23%,平均检测速率是Faster R-CNN模型的9.6倍,所提方法在红外弱光环境下能够更加快速准确地检测出铁路异物,主客观评价均优于对比方法。  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1001-5965  
入藏号:  
CSCD:7543864  
Indexed Date:  
2023-12-08  
Institution Address:  
兰州交通大学电子与信息工程学院;;甘肃省人工智能与图形图像处理工程研究中心, ;;甘肃省人工智能与图形图像处理工程研究中心, 兰州;;兰州, ;; 730070;;730070  
Institution Address:  
兰州交通大学电子与信息工程学院, 兰州, 甘肃 730070, 中国  
Institution Address:  
School of Electronic and Information Engineering, Lanzhou Jiaotong University;;Gansu Artificial Intelligence and Graphics and Image Processing Engineering Research Center, ;;Gansu Artificial Intelligence and Graphics and Image Processing Engineering Research Center, Lanzhou;;Lanzhou, ;; 730070;;730070  
Institution Address:  
School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China  
Institution:  
兰州交通大学电子与信息工程学院;;甘肃省人工智能与图形图像处理工程研究中心  
Institution:  
兰州交通大学电子与信息工程学院  
Institution:  
School of Electronic and Information Engineering, Lanzhou Jiaotong University;;Gansu Artificial Intelligence and Graphics and Image Processing Engineering Research Center  
Institution:  
School of Electronic and Information Engineering, Lanzhou Jiaotong University

### 382. MSA R-CNN: A comprehensive approach to remote sensing object detection

摘要: Remote sensing scene understanding is crucial for extracting valuable information from high-resolution images, including object detection and classification. Traditional object detection methods face challenges in handling the diverse scales, orientations, and complex backgrounds present in remote sensing data. In this paper, we propose a novel remote sensing scene understanding system called multiscale-attention R-CNN (MSA R-CNN), which incorporates a super multiscale feature extraction network (SMENet) for enhanced feature extraction from multiscale images, an adaptive dynamic inner lateral (ADIL) connection module to tackle information loss in feature pyramid networks (FPN), and a distributed lightweight attention module (DLAM) to refine feature information processing. Furthermore, a new dataset combining the DIOR and DOTA datasets is introduced to extract the background information of detected objects and evaluate the proposed system's performance. MSA R-CNN achieved an mAP of 74.37% on the DIOR dataset when the gamma value was set to 0.2 and 81.97% on the DOTA dataset when the gamma value was set to 0.1 with the same learning rate, outperforming state-of-the-art models on both datasets. The proposed system demonstrates significant improvements in both object detection and background information extraction, providing a comprehensive solution for remote sensing scene understanding.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
29  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
29  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001133018300001  
Indexed Date:  
2024-01-11  
Institution Address:  
Sejong Univ, Dept Intelligent Mechatron Engn, Seoul 05006, South Korea  
Institution Address:  
Sejong Univ, Dept Comp Sci & Engn, Seoul 05006, South Korea  
Institution Address:  
Southwest Jiaotong Univ, Dept Geosci & Environm Engn, Chengdu 610097, Peoples R China  
Institution:  
Institution:  
Institution:

### 383. DPCTN: Dual path context-aware transformer network for medical image

摘要: Accurate segmentation of lesions in medical images is a key step to assist clinicians in diagnosis and analysis. Most studies combine the Transformer model with CNN at a single scale or use the highest-level feature tensor extracted by CNN as input to Transformer without fully exploiting Transformer's potential. In addition, for the problems of structural boundary blurring, heterogeneous textures, etc., in medical images, most existing methods pay attention to using contour information to solve this problem but simply fuse the contour information and ignore the potential relationship between the regions and the contours. We propose the DPCTN network based on the traditional encoding-decoding structure, consisting of the CNN, Transformer dual backbone networks and parallel attention mechanisms, to achieve accurate segmentation in medical image lesions. Local and global multiscale feature information is extracted by CNN and Transformer. The Transformer block of channel cross fusion can implement multiscale information fusion of the high-level local features and reduce the impact of the redundant information. The dual backbone feature fusion module effectively couples the local and global high-level feature information. The decoder refines and enriches the boundary and regional features, layer by layer, to achieve effective supervision of the boundary and region. Considering the possible dimension collapse in the attention mechanism, a novel three branch transposed self-attention module is designed to reduce the information loss caused by feature pooling. To verify the effectiveness of our proposed method, subjective and objective comparative experiments and ablation experiments were performed on four medical segmentation tasks, polyps, skin lesions, glands and breast tumors. A large number of experimental results show that our method is superior to the current state-of-the-art method, reduces the standard deviation and is more robust. Source code is released at https://github.com/sd-spf/DPCTN.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
17  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
17  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001025808200001  
Indexed Date:  
2023-07-22  
Institution Address:  
Shandong Technol & Business Univ, Sch Comp Sci & Technol, Yantai, Peoples R China  
Institution Address:  
Shandong Technol & Business Univ, Coinnovat Ctr Shandong Coll & Univ Future Intellig, Yantai, Peoples R China  
Institution:  
Institution:

### 384. Self-supervised learning-based Multi-Scale feature Fusion Network for

摘要: Understanding prognosis and mortality is critical for evaluating the treatment plan of patients. Advances in digital pathology and deep learning techniques have made it practical to perform survival analysis in whole slide images (WSIs). Current methods are usually based on a multi-stage framework which includes patch sampling, feature extraction and prediction. However, the random patch sampling strategy is highly unstable and prone to sampling non-ROI. Feature extraction typically relies on hand-crafted features or convolutional neural networks (CNNs) pre-trained on ImageNet, while the artificial error or domain gaps may affect the survival prediction performance. Besides, the limited information representation of local sampling patches will create a bottleneck limitation on the effectiveness of prediction. To address the above challenges, we propose a novel patch sampling strategy based on image information entropy and construct a Multi-Scale feature Fusion Network (MSFN) based on self-supervised feature extractor. Specifically, we adopt image information entropy as a criterion to select representative sampling patches, thereby avoiding the noise interference caused by random to blank regions. Meanwhile, we pretrain the feature extractor utilizing self -supervised learning mechanism to improve the efficiency of feature extraction. Furthermore, a global-local feature fusion prediction network based on the attention mechanism is constructed to improve the survival prediction effect of WSIs with comprehensive multi-scale information representation. The proposed method is validated by adequate experiments and achieves competitive results on both of the most popular WSIs survival analysis datasets, TCGA-GBM and TCGA-LUSC. Code and trained models are made available at: https://github.com/Mercuriiio/MSFN.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
8  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:000923216300001  
PubMed ID:  
36586231  
Indexed Date:  
2023-02-20  
Institution Address:  
Macau Univ Sci & Technol, Fac Innovat Engn, Macau 999078, Peoples R China  
Institution Address:  
Peng Cheng Lab, Shenzhen 518055, Peoples R China  
Institution Address:  
China Univ Petr, Coll Comp Sci & Technol, Qingdao 266580, Peoples R China  
Institution:  
Institution:  
Institution:

### 385. A multi-view convolutional neural network based on cross-connection and

摘要: Convolutional neural networks are widely used in image feature extraction, but the architecture of existing models is overly complex. To solve this problem, this paper proposes a novel convolutional neural network named CFRW, which consists of two parts, C-FnetT and R-WN. C-FnetT focuses on deepening the network through the Cross-Connection algorithm to extract high-level image features. R-WN extracts rich multiscale feature information by the Residual-Wider module with parallel branches. At the same time, a new fusion algorithm, FA, is designed, which maps multidimensional features to a single-dimensional feature space and assigns weight to feature information from the channel dimension. By learning multiangle information of image features from the local space to the global space, the feature expression ability of the convolutional neural network is enhanced. C-FnetT and R-WN are fused by FA. Through a series of comparative experiments on public datasets MNIST, FashionMNIST, SVHN, CIFAR10, CIFAR100 and Oxford Flowers 17, the results show that the designed CFRW with a multirepresentation structure and fewer layers achieves promising performance in image recognition compared with other methods.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000871297400002  
Indexed Date:  
2022-10-31  
Institution Address:  
North China Elect Power Univ, Sch Control & Comp Engn, Beijing 102206, Peoples R China  
Institution:

### 386. RCCT-ASPPNet: Dual-Encoder Remote Image Segmentation Based on

摘要: Remote image semantic segmentation technology is one of the core research elements in the field of computer vision and has a wide range of applications in production life. Most remote image semantic segmentation methods are based on CNN. Recently, Transformer provided a view of long-distance dependencies in images. In this paper, we propose RCCT-ASPPNet, which includes the dual-encoder structure of Residual Multiscale Channel Cross-Fusion with Transformer (RCCT) and Atrous Spatial Pyramid Pooling (ASPP). RCCT uses Transformer to cross fuse global multiscale semantic information; the residual structure is then used to connect the inputs and outputs. ASPP based on CNN extracts contextual information of high-level semantics from different perspectives and uses Convolutional Block Attention Module (CBAM) to extract spatial and channel information, which will further improve the model segmentation ability. The experimental results show that the mIoU of our method is 94.14% and 61.30% on the datasets Farmland and AeroScapes, respectively, and that the mPA is 97.12% and 84.36%, respectively, both outperforming DeepLabV3+ and UCTransNet.  
Web of Science 核心合集中的 "被引频次":  
18  
在中国科学引文数据库中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
20  
eISSN:  
2072-4292  
入藏号:  
WOS:000916422000001  
Indexed Date:  
2023-02-09  
Institution Address:  
Anhui Univ, Natl Engn Res Ctr Anal & Applicat Agroecol Big Dat, Hefei 230601, Peoples R China  
Institution Address:  
Anhui Univ, Sch Internet, Hefei 230039, Peoples R China  
Institution:  
Institution:

### 387. Detection of Residual Film on the Field Surface Based on Faster R-CNN

摘要: After the residual film recycling machine recovers the film, some small pieces of the film will remain on the surface of the field. To solve the problem of collecting small pieces of film, it is necessary to develop a piece of intelligent picking equipment. The detection of small pieces of film is the first problem to be solved. This study proposes a method of an object detection algorithm fusing multi-scale features (MFFM Faster R-CNN) based on improved Faster R-CNN. Based on the Faster R-CNN model, the feature pyramid network is added to solve the problem of multiscale change of residual film. The convolution block attention module is introduced to enhance the feature extraction ability of the model. The Soft-NMS algorithm is used instead of the NMS algorithm to improve the detection accuracy of the model in the RPN network. The experimental results show that the model is able to effectively detect surface residual film in complex environments, with an AP of 83.45%, F1-score of 0.89, and average detection time of 248.36 ms. The model is compared with SSD and YOLOv5 under the same experimental environment and parameters, and it is found that the model not only ensures high-precision detection but also ensures real-time detection. This lays the theoretical foundation for the subsequent development of field surface residual film intelligent picking equipment.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
eISSN:  
2077-0472  
入藏号:  
WOS:001016862700001  
Indexed Date:  
2023-07-07  
Institution Address:  
Xinjiang Agr Univ, Coll Mech & Elect Engn, Urumqi 830052, Peoples R China  
Institution Address:  
Xinjiang Acad Agr Sci, Res Inst Agr Mechanizat, Urumqi 830091, Peoples R China  
Institution:  
Institution:

### 388. Multi-View Scene Classification Based on Feature Integration and

摘要: Leveraging multi-view remote sensing images in scene classification tasks significantly enhances the precision of such classifications. This approach, however, poses challenges due to the simultaneous use of multi-view images, which often leads to a misalignment between the visual content and semantic labels, thus complicating the classification process. In addition, as the number of image viewpoints increases, the quality problem for remote sensing images further limits the effectiveness of multi-view image classification. Traditional scene classification methods predominantly employ SoftMax deep learning techniques, which lack the capability to assess the quality of remote sensing images or to provide explicit explanations for the network's predictive outcomes. To address these issues, this paper introduces a novel end-to-end multi-view decision fusion network specifically designed for remote sensing scene classification. The network integrates information from multi-view remote sensing images under the guidance of image credibility and uncertainty, and when the multi-view image fusion process encounters conflicts, it greatly alleviates the conflicts and provides more reasonable and credible predictions for the multi-view scene classification results. Initially, multi-scale features are extracted from the multi-view images using convolutional neural networks (CNNs). Following this, an asymptotic adaptive feature fusion module (AAFFM) is constructed to gradually integrate these multi-scale features. An adaptive spatial fusion method is then applied to assign different spatial weights to the multi-scale feature maps, thereby significantly enhancing the model's feature discrimination capability. Finally, an evidence decision fusion module (EDFM), utilizing evidence theory and the Dirichlet distribution, is developed. This module quantitatively assesses the uncertainty in the multi-perspective image classification process. Through the fusing of multi-perspective remote sensing image information in this module, a rational explanation for the prediction results is provided. The efficacy of the proposed method was validated through experiments conducted on the AiRound and CV-BrCT datasets. The results show that our method not only improves single-view scene classification results but also advances multi-view remote sensing scene classification results by accurately characterizing the scene and mitigating the conflicting nature of the fusion process.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
eISSN:  
2072-4292  
入藏号:  
WOS:001183133400001  
Indexed Date:  
2024-03-30  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Remote Sensing & Geomat Engn, Nanjing 210044, Peoples R China  
Institution Address:  
Beijing Normal Univ, State Key Lab Remote Sensing Sci, Beijing 100875, Peoples R China  
Institution Address:  
Emory Univ, Dept Environm Sci, Atlanta, GA 30322 USA  
Institution:  
Institution:  
Institution:

### 389. Improved Cascade RCNN algorithm for character detection of cast pipe

摘要: Due to the issues of blurred background, small character areas, inconsistent engraving positions and paint occlusion in the images of cast pipe characters collected from industrial sites, the existing models struggle to meet the detection accuracy requirements of industrial environments. To address these issues, an improved Cascade Region Convolutional Neural Network (RCNN)for cast pipe character detection was proposed. The improvements were made to the feature pyramid by introducing the Small Target Enhancement Feature Pyramid Networks (STE-FPN), which utilized the feature enhancement capability of multi-scale feature fusion to enrich the feature information of small target characters on cast pipes. The Self-Calibrated/Pooling ResNeSt (SCP-ResNeSt) was introduced as the feature extraction network, utilizing self-calibrated convolutions and pooling operations to enhance the efficiency of extracting complex background features of cast pipe characters. The improvements were made to the cascade structure by introducing a Mask branch structure, which adaptively detected character areas and removes interference regions,thus optimizing the detection results. The improved algorithm was tested on the cast pipe dataset,and mean Average Precision (mAP) was 99.1%,which increased the precision by 2.3% compared to the Cascade RCNN. The precision showed that the improved performance was superior to the original algorithm.  
摘要:  
由于工业现场采集的铸管字符图像存在背景模糊、字符区域占比小、刻字位置不固定、油漆遮挡等问题,导致现有模型的检测精度难以满足工业现场的需求。针对上述问题,提出改进的Cascade RCNN铸管字符检测算法。首先对特征金字塔进行改进,提出融合小目标增强的特征金字塔(STE-FPN),利用多尺度特征融合的特征增强能力丰富铸管小目标字符的特征信息。其次引入自矫正/池化的ResNeSt(SCP-ResNeSt)作为特征提取网络,利用自矫正卷积和池化操作以提升背景复杂的铸管字符特征提取效率。最后对级联结构进行改进,引进Mask分支结构,可以自适应地检测字符区域并去除干扰区域,优化了检测结果。将改进后的算法在铸管数据集上进行测试,其平均检测精度mAP为99.1%,比原Cascade RCNN算法提高了2.3%,得到的精度表明改进后的性能优于原算法。  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1006-5911  
入藏号:  
CSCD:7851669  
Indexed Date:  
2025-02-09  
Institution Address:  
中国科学院网络化控制系统重点实验室;;中国科学院沈阳自动化研究所;;中国科学院机器人与智能制造创新研究院, 中国科学院网络化控制系统重点实验室;;;;, 沈阳;;沈阳;;沈阳, 辽宁;;辽宁;;辽宁 110016;;110016;;110169, 中国  
Institution Address:  
沈阳工业大学软件学院, 沈阳, 辽宁 110870, 中国  
Institution Address:  
上海波士内智能科技有限公司, 上海 201612, 中国  
Institution Address:  
沈阳中科博微科技股份有限公司, 沈阳, 辽宁 110180, 中国  
Institution Address:  
Key Laboratory of Networked Control Systems,Chinese Academy of Sciences;;Shenyang Institute of Automation,Chinese Academy of Sciences;;Institutes for Robotics and Intelligent Manufacturing,Chinese Academy of Sciences, Key Laboratory of Networked Control Systems,Chinese Academy of Sciences;;;;, Shenyang;;Shenyang;;Shenyang, ;;;; 110016;;110016;;110169  
Institution Address:  
School of Software,Shenyang University of Technology, Shenyang, Liaoning 110870, China  
Institution Address:  
Shanghai Ballsnow Intelligent Technology Co.,Ltd., Shanghai 201612, China  
Institution Address:  
Shenyang Zhongke Microcyber Technology Co.,Ltd., Shenyang, Liaoning 110180, China  
Institution:  
中国科学院网络化控制系统重点实验室;;中国科学院沈阳自动化研究所;;中国科学院机器人与智能制造创新研究院  
Institution:  
沈阳工业大学软件学院  
Institution:  
上海波士内智能科技有限公司  
Institution:  
沈阳中科博微科技股份有限公司  
Institution:  
Key Laboratory of Networked Control Systems,Chinese Academy of Sciences;;Shenyang Institute of Automation,Chinese Academy of Sciences;;Institutes for Robotics and Intelligent Manufacturing,Chinese Academy of Sciences  
Institution:  
School of Software,Shenyang University of Technology  
Institution:  
Shanghai Ballsnow Intelligent Technology Co.,Ltd.  
Institution:  
Shenyang Zhongke Microcyber Technology Co.,Ltd.

### 390. Dynamic Auxiliary Soft Labels for decoupled learning

摘要: The long-tailed distribution in the dataset is one of the major challenges of deep learning. Convolutional Neural Networks have poor performance in identifying classes with only a few samples. For this problem, it has been proved that separating the feature learning stage and the classifier learning stage improves the performance of models effectively, which is called decoupled learning. We use soft labels to improve the performance of the decoupled learning framework by proposing a Dynamic Auxiliary Soft Labels (DaSL) method. Specifically, we design a dedicated auxiliary network to generate auxiliary soft labels for the two different training stages. In the feature learning stage, it helps to learn features with smaller variance within the class, and in the classifier learning stage it helps to alleviate the overconfidence of the model prediction. We also introduce a feature-level distillation method for the feature learning, and improve the learning of general features through multi-scale feature fusion. We conduct extensive experiments on three long-tailed recognition benchmark datasets to demonstrate the effectiveness of our DaSL.(C) 2022 Elsevier Ltd. All rights reserved.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0893-6080  
eISSN:  
1879-2782  
入藏号:  
WOS:000792692200008  
PubMed ID:  
35421708  
Indexed Date:  
2022-05-23  
Institution Address:  
Nanjing Univ, State Key Lab Novel Software Technol, Nanjing, Peoples R China  
Institution Address:  
Nanjing Univ, Sch Artificial Intelligence, Nanjing, Peoples R China  
Institution Address:  
Nanjing Univ, Dept Comp Sci & Technol, Nanjing, Peoples R China  
Institution Address:  
Nanjing Univ, Sch Elect Sci & Engn, Nanjing, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 391. CIFTC-Net: Cross information fusion network with transformer and CNN for

摘要: Polyp segmentation plays a crucial role in the early diagnosis and treatment of colorectal cancer, which is the third most common cancer worldwide. Despite remarkable successes achieved by recent deep learning-related works, accurate segmentation of polyps remains challenging due to the diversity in their shapes, sizes, appearances, and other factors. To address these problems, a novel cross information fusion network with Transformer and convolutional neural network (CNN) for polyp segmentation, named CIFTC-Net, is proposed to improve the segmentation performance of colon polyps. In particular, a dual-branch encoder with Pyramid Vision Transformer (PVT) and ResNet50 is employed to take full advantage of both the global semantic information and local spatial features to enhance the feature representation ability. To effectively fuse the two types of features, a new global-local feature fusion (GLFF) module is designed. Additionally, in the PVT branch, a multi-scale feature integration (MSFI) module is introduced to fuse multi-scale features adaptively. At the bottom of the model, a multi-scale atrous pyramid bridging (MSAPB) module is proposed to achieve rich and robust multi-level features and improve the segmentation accuracy. Experimental results on four public polyp segmentation datasets demonstrate that CIFTC-Net surpasses current state-of-the-art methods across various metrics, showcasing its superiority in segmentation accuracy, generalization ability, and handling of complex images.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0141-9382  
eISSN:  
1872-7387  
入藏号:  
WOS:001351510900001  
Indexed Date:  
2024-11-16  
Institution Address:  
Univ Shanghai Sci & Technol, Sch Hlth Sci & Engn, Shanghai, Peoples R China  
Institution Address:  
Shanghai Univ Med & Hlth Sci, Sch Med Instruments, Shanghai, Peoples R China  
Institution Address:  
Shanghai Guanghua Hosp Integrated Tradit Chinese &, Gastroenterol Dept, Shanghai, Peoples R China  
Institution Address:  
Tongji Univ, Tongji Hosp, Equipment & Mat Dept, Shanghai, Peoples R China  
Institution:  
Institution:  
Institution:  
Shanghai Guanghua Hosp Integrated Tradit Chinese &  
Institution:

### 392. Global attention mechanism with real-time semantic segmentation network

摘要: The lightweight network structure cannot sufficiently extract effective semantic information from feature maps,and the unreasonable design of the semantic information and spatial detail information fusion block leads to a decrease in segmentation accuracy.To address these problems,a global attention mechanism with a real-time semantic segmentation network (GaSeNet) is proposed in the paper.First,a global attention mechanism is introduced into the semantic branch of the dual-branch structure.The convolutional neural network is then guided in the two dimensions of channel and space to focus on the semantic categories related to the segmentation task to extract remarkably effective semantic information.Second,a mixed hole convolution block is designed in the spatial detail branch,and the receptive field is enlarged while maintaining the size of the convolution kernel to obtain additional global spatial detail information and compensate for the loss of key feature information.The feature fusion module is then redesigned,and the deep aggregation pyramid pooling module is introduced to fuse feature maps of different scales comprehensively,thereby improving the semantic segmentation performance of the network.Finally,the proposed method is tested on CamVid and Vaihingen datasets.Compared with the latest semantic segmentation algorithm,GaSeNet improves the segmentation accuracy by 4.29% and 16.06%.Experimental results verify the effectiveness of this method in dealing with real-time semantic segmentation problems.  
摘要:  
针对轻量化网络结构从特征图提取有效语义信息不足,以及语义信息与空间细节信息融合模块设计不合理而导致分割精度降低的问题,本文提出一种结合全局注意力机制的实时语义分割网络(global attention mechanism with real time semantic segmentation network,GaSeNet)。首先在双分支结构的语义分支中引入全局注意力机制,在通道与空间两个维度引导卷积神经网来关注与分割任务相关的语义类别,以提取更多有效语义信息;其次在空间细节分支设计混合空洞卷积块,在卷积核大小不变的情况下扩大感受野,以获取更多全局空间细节信息,弥补关键特征信息损失。然后重新设计特征融合模块,引入深度聚合金塔池化,将不同尺度的特征图深度融合,从而提高网络的语义分割性能。最后将所提出的方法在CamVid数据集和Vaihingen数据集上进行实验,通过与最新的语义分割方法对比分析可知,GaSeNet在分割精度上分别提高了4.29%、16.06%,实验结果验证了本文方法处理实时语义分割问题的有效性。  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1673-4785  
入藏号:  
CSCD:7431775  
Indexed Date:  
2023-07-01  
Institution Address:  
河南师范大学计算机与信息工程学院;;智慧商务与物联网技术河南省工程实验室, ;;智慧商务与物联网技术河南省工程实验室, 新乡;;新乡, 河南;;河南 453007;;453007, 中国  
Institution Address:  
河南师范大学软件学院, 新乡, 河南 453007, 中国  
Institution Address:  
中国人民公安大学国家安全学院, 北京 100038, 中国  
Institution Address:  
河南师范大学计算机与信息工程学院, 新乡, 河南 453007, 中国  
Institution Address:  
College of Computer and Information Engineering,He'nan Normal University;;Engineering Lab of He'nan Province for Intelligence Business &Internet of Things, ;;Engineering Lab of He'nan Province for Intelligence Business &Internet of Things, Xinxiang;;Xinxiang, ;; 453007;;453007  
Institution Address:  
College of Software,He'nan Normal University, Xinxiang, Henan 453007, China  
Institution Address:  
National Security Academy,People's Public Security University of China, Beijing 100038, China  
Institution Address:  
College of Computer and Information Engineering,He'nan Normal University, Xinxiang, Henan 453007, China  
Institution:  
河南师范大学计算机与信息工程学院;;智慧商务与物联网技术河南省工程实验室  
Institution:  
河南师范大学软件学院  
Institution:  
中国人民公安大学国家安全学院  
Institution:  
河南师范大学计算机与信息工程学院  
Institution:  
College of Computer and Information Engineering,He'nan Normal University;;Engineering Lab of He'nan Province for Intelligence Business &Internet of Things  
Institution:  
College of Software,He'nan Normal University  
Institution:  
National Security Academy,People's Public Security University of China  
Institution:  
College of Computer and Information Engineering,He'nan Normal University

### 393. An effective multi-scale interactive fusion network with hybrid

摘要: Smoke has visually elusive appearances, especially in low-light conditions, so it is quite difficult to quickly and accurately detect smoke from images. To address these challenges, we design a dual-encoder structure of Transformer and Convolutional Neural Network (CNN) to propose an effective Multi-scale Interactive Fusion Network (MIFNet) for smoke image segmentation. To improve the presentation of features, we propose a Local Feature Enhancement Propagation (LFEP) module to enhance spatial details. To optimize global and local features for efficient fusion, we integrate LFEP into the original Transformer to replace the traditional multi-head self-attention mechanism. Then, we propose a Multi-level Attention Coupled Module (MACM) to fuse Transformer and CNN features of the dual-encoder. MACM can flexibly focus on information interaction between different levels of two encoding paths. Finally, we design a Prior-guided Multi-scale Fusion Decoder (PMFD), which combines prior knowledge with a multi-scale feature fusion strategy to improve the performance of segmentation. Experimental results demonstrate that MIFNet substantially outperforms the state-of-the-art methods. MIFNet achieves a mean Intersection over Union (mIoU) of 81.6 % on the synthetic smoke (SYN70 K) dataset, and a remarkable accuracy of 98.3 % on the forest smoke dataset.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:001413917300001  
Indexed Date:  
2025-02-12  
Institution Address:  
Shanghai Normal Univ, Coll Math & Sci, Shanghai 200234, Peoples R China  
Institution Address:  
Shanghai Normal Univ, Coll Informat Mech & Elect Engn, Shanghai 201418, Peoples R China  
Institution:  
Institution:

### 394. ATNet: A Defect Detection Framework for X-ray Images of DIP Chip Lead

摘要: In order to improve the production quality and qualification rate of chips, X-ray nondestructive imaging technology has been widely used in the detection of chip defects, which represents an important part of the quality inspection of products after packaging. However, the current traditional defect detection algorithm cannot meet the demands of high accuracy, fast speed, and real-time chip defect detection in industrial production. Therefore, this paper proposes a new multi-scale feature fusion module (ATSPPF) based on convolutional neural networks, which can more fully extract semantic information at different scales. In addition, based on this module, we design a deep learning model (ATNet) for detecting lead defects in chips. The experimental results show that at 8.2 giga floating point operations (GFLOPs) and 146 frames per second (FPS), mAP(0.5) and mAP(0.5-0.95) can achieve an average accuracy of 99.4% and 69.3%, respectively, while the detection speed is faster than the baseline yolov5s by nearly 50%.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
eISSN:  
2072-666X  
入藏号:  
WOS:001036725000001  
PubMed ID:  
37512688  
Indexed Date:  
2023-08-10  
Institution Address:  
Guangdong Univ Technol, Sch Mech & Elect Engn, Guangzhou 510006, Peoples R China  
Institution Address:  
Guangdong Univ Technol, Sch Informat Engn, Guangzhou 510006, Peoples R China  
Institution Address:  
China Elect Prod Reliabil & Environm Testing Res I, Guangzhou 511370, Peoples R China  
Institution:  
Institution:  
Institution:  
China Elect Prod Reliabil & Environm Testing Res I

### 395. Few-shot object detection via fusing multi-scale and attention mechanism

摘要: The existing two-stage few-shot object detection methods based on fine-tuning are not sensitive to the features of new classes, which will cause misjudgment of new classes into base classes with high similarity to them, thus affecting the detection performance of the model. To address the above issue, a few-shot object detection algorithm that incorporates multiscale and attention mechanism was proposed, namely MA-FSOD (Few-Shot Object Detection via fusing Multi-scale and Attention mechanism). Firstly, grouped convolutions and large convolution kernels were used to extract more classdiscriminative features in the backbone network, and Convolutional Block Attention Module (CBAM) was added to achieve adaptive feature augmentation. Then, a modified pyramid network was used to achieve multi-scale feature fusion, which enables Region Proposal Network (RPN) to accurately find Regions of Interest (RoI) and provide more abundant highquality positive samples from multiple scales to the classification head. Finally, the cosine classification head was used for classification in the fine-tuning stage to reduce the intra-class variance. Compared with the Few-Shot object detection via Contrastive proposal Encoding (FSCE) algorithm on PASCAL-VOC 2007/2012 dataset, the MA-FSOD algorithm improved AP\_(50) for new classes by 5.6 percentage points; and on the more challenging MSCOCO dataset, compared with Meta-Faster- RCNN, the APs corresponding to 10-shot and 30-shot were improved by 0.1 percentage points and 1.6 percentage points, respectively. Experimental results show that MA-FSOD can more effectively alleviate the misclassification problem and achieve higher accuracy in few-shot object detection than some mainstream few-shot object detection algorithms.  
摘要:  
现有基于微调的二阶段小样本目标检测方法对新类特征不敏感,易将新类别误判成与它相似度高的基类,影响模型的检测性能。针对上述问题,提出一种融合多尺度和注意力机制的小样本目标检测(MA-FSOD)算法。首先在骨干网络使用分组卷积和大卷积核提取更具类别区分性的特征,并加入卷积注意力模块(CBAM)实现特征的自适应增强;再通过改进的金字塔网络实现多尺度的特征融合,使候选框生成网络(RPN)可以准确找到感兴趣区域(RoI),从多个尺度向分类头提供更丰富的高质量正样本;最后在微调阶段采用余弦分类头进行分类,降低类内方差。在PASCAL-VOC 2007/2012数据集上与基于候选框编码对比损失的小样本目标检测(FSCE)算法相比,MA-FSOD算法对新类的AP\_(50)提升了5.6个百分点;在更具挑战性的MSCOCO数据集中,与Meta-Faster-RCNN相比,10-shot和30-shot对应的AP则分别提升了0.1个百分点和1.6个百分点。实验结果表明,相较于一些主流的小样本目标检测算法, MA-FSOD算法能更有效地缓解误分类问题,实现更高精度的小样本目标检测。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1001-9081  
入藏号:  
CSCD:7731731  
Indexed Date:  
2024-08-08  
Institution Address:  
北京联合大学, 北京市信息服务工程重点实验室, 北京 100101, 中国  
Institution Address:  
北京联合大学;;北京开放大学科技学院, 北京市信息服务工程重点实验室;;, ;;, 北京;;北京 100101;;100081, 中国  
Institution Address:  
Beijing Union University, Beijing Key Laboratory of Information Service Engineering, Beijing 100101, China  
Institution Address:  
Beijing Union University;;College of Science and Technology, Beijing Open University, Beijing Key Laboratory of Information Service Engineering;;, ;;, Beijing;;Beijing 100101;;100081  
Institution:  
北京联合大学  
Institution:  
北京联合大学;;北京开放大学科技学院  
Institution:  
Beijing Union University  
Institution:  
Beijing Union University;;College of Science and Technology, Beijing Open University

### 396. RMCA U-net: Hard exudates segmentation for retinal fundus images

摘要: Hard exudate plays an important role in grading diabetic retinopathy (DR) as a critical indicator. Therefore, the accurate segmentation of hard exudates is of clinical importance. However, the percentage of hard exudates in the whole fundus image is relatively small, and their shapes are often irregular and the contrasts are usually not high enough. Hence, they are prone to misclassifications e.g., misclassified as part of the optic disc structure or cotton wool spots, which results in the low segmentation accuracy and efficiency. This paper proposes a novel neural network RMCA U-net to accurately segmentation hard exudate in fundus images. The network features a U-shape framework combined with a residual structure to obtain the subtle features of hard exudate. A multi-scale feature fusion (MSFF) module and an improved channel attention (CA) module are designed and involved to effectively segmentation sparse small lesions. The proposed method in this paper has been trained and evaluated on three data sets: IDRID, Kaggle and one local data set. Experiments are shown and indicate that RMCA U-net of this paper is superior to the other convolutional neural networks. The method in this paper is increased by 6% higher in PR-MAP than U-net on the IDRID dataset, increased by 10% in Recall than U-net on the Kaggle dataset and increased by 20% in F1-score than U-net on the local dataset.  
Web of Science 核心合集中的 "被引频次":  
13  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
14  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001048783000001  
Indexed Date:  
2023-08-23  
Institution Address:  
Univ Shanghai Sci & Technol, Sch Opt Elect & Comp Engn, Shanghai, Peoples R China  
Institution Address:  
UCL, Inst Hlth Informat, London, England  
Institution Address:  
Tongji Univ, Shanghai Inst Intelligent Sci & Technol, Shanghai, Peoples R China  
Institution:  
Institution:  
Institution:

### 397. Transformer-based multi-scale feature fusion network for remote sensing

摘要: Change detection (CD) is the operation of quantitatively analyzing the surface changes of a phenomenon or objects over two different times. Lately, CD based on deep learning has developed to become more and more powerful, and convolutional neural networks (CNNs) have dominated the field of remote sensing (RS) CD. In particular, in many fields of computer vision, neural networks based on U-Net network and skip connections have been generally used. However, despite the excellent performance achieved by CNN, it does not learn global and longrange semantic information interaction well due to the locality of convolutional operations. The recently proposed Swin-UNet in the field of medical image segmentation achieved excellent results, which is a U-Net-like pure transformer. In the face of the challenge of segmentation accuracy, the Swin transformer has demonstrated strong capabilities. The Swin transformer block (STB) consists of residual connected STBs used in SwinIR to enhanced training stability. We began to try to incorporate them into our network for RS CD. Finally, we propose a transformer-based multi-scale feature fusion model (TMFF), including decoder, encoder, and skip connection structure, for RS image CD. We modify the original U-Net architecture so that it can better aggregate semantic features at all levels. Our proposed TMFF achieves impressive results through experiments on three datasets; (c) 2022 Society of Photo-Optical Instrumentation Engineers (SPIE)  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
eISSN:  
1931-3195  
入藏号:  
WOS:000917345800039  
Indexed Date:  
2023-02-17  
Institution Address:  
Shandong Technol & Business Univ, Sch Informat & Elect Engn, Yantai, Peoples R China  
Institution Address:  
Shandong Technol & Business Univ, Sch Comp Sci & Technol, Yantai, Peoples R China  
Institution:  
Institution:

### 398. Attention mechanism combined with residual recurrent neural network for

摘要: In the task of sound event detection and localization (SEDL) in a complex environment, the acoustic signals of different events usually have nonlinear superposition, so the detection and localization effect is not good. Given this, this paper is based on the Residual-spatially and channel Squeeze-Excitation (Res-scSE) model. Combined with Multiple-scale Convolutional Recurrent Neural Network (M-CRNN), the Res-scSE-CRNN model is proposed. Firstly, to solve the problem of insufficient extraction of time-frequency feature in single-size convolution kernel, multi-scale feature fusion is carried out by using the feature hierarchy of the convolutional neural network to improve the accuracy of detection. Secondly, aiming at the problem of overlapping audio event localization accuracy is not high, with Res-scSE to replace common convolution module and add residual structure to strengthen the feature extraction, and combining with an attention mechanism to enhance neural network channels and spatial relationships, to improve the network to extract the characteristics of directivity, achieve the goal of the overlapped audio localization. In this paper, experiments are carried out in the open dataset DCASE2019, and evaluation indicators are used to analyze the effectiveness of the proposed model and baseline model in the detection and localization of audio events. The results show that compared with the M-CRNN model, the detection error rate of Res-scSE-CRNN model is reduced 4%, the F1-Score is increased 3.4%, the localization error is reduced by 22.8 & DEG;, and the frame recall rate is increased 3%.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
1687-4722  
入藏号:  
WOS:000893897600001  
Indexed Date:  
2022-12-16  
Institution Address:  
Harbin Univ Sci & Technol, Dept Sch Measurement & Commun Engn, Harbin 150080, Peoples R China  
Institution Address:  
Beidahuang Ind Grp Gen Hosp, Harbin 150088, Peoples R China  
Institution Address:  
Hainan Univ, Mech & Elect Engn Coll, Haikou 570228, Peoples R China  
Institution Address:  
Guangzhou Univ, Sch Elect & Commun Engn, Guangzhou 510006, Peoples R China  
Institution:  
Institution:  
Beidahuang Ind Grp Gen Hosp  
Institution:  
Institution:

### 399. Intelligent Fault Diagnosis Method of Wind Turbines Planetary Gearboxes

摘要: Due to the powerful capability of feature extraction, convolutional neural network (CNN) is increasingly applied to the fault diagnosis of key components of rotating machineries. Due to the shortcomings of traditional CNN-based fault diagnosis methods, the continuous convolution and pooling operations result in the constant decrease of feature resolution, which may cause the loss of some subtle fault information in the samples. This paper proposes a CNN-based model with improved structure multi-scale dense fusion network (MSDFN) to realize the fault diagnosis of wind turbines planetary gearboxes under complicated working conditions. First, the continuous wavelet transform is applied to preprocess the vibration signals, and the two-dimensional wavelet time-frequency diagrams are used as the network input. Then, the multi-scale feature fusion (MSFF) module and a feature of maximum (FoM) module are used in the extraction and classification stages of fault features, respectively. Next, the multi-scale features of each network layer are fused to enhance the fault features. Finally, the high fault diagnosis accuracy is achieved by extracting the separable fusion result of fault features. The proposed method achieves more than 99% fault diagnosis average accuracy on a planetary gearbox dataset. The comparative experimental results verify the effectiveness of the proposed method and its superiority to some mainstream approaches. The ablation study further confirms that MSFF module and FoM module play the positive role in fault diagnosis.  
Web of Science 核心合集中的 "被引频次":  
6  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
2296-598X  
入藏号:  
WOS:000731702500001  
Indexed Date:  
2021-12-27  
Institution Address:  
Chongqing Univ, Coll Automat, Chongqing, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Coll Automat, Chongqing, Peoples R China  
Institution:  
Institution:

### 400. Cross-Scale Feature Interaction Network for Semantic Segmentation in

摘要: Semantic Segmentation in side-scan sonar images (SSS-Seg) is an emerging topic and plays an important function in sonar image interpretation. However, due to the interference of seabed reverberation noise, complex background information, and the unique characteristics of sonar images, the direct application of natural scene image semantic segmentation methods to SSS-Seg fails to achieve satisfactory results. For SSS image semantic segmentation, existing challenges include the inability to effectively distinguish between similar objects, sensitivity to noise, and loss of critical feature details during segmentation. In this article, we propose a novel cross-scale feature interaction network (CSFINet) to address these challenges and achieve semantic segmentation for different underwater objects in SSS images. Specifically, the cross-scale feature selection module filters spatial detail features and abstracts semantic information. The multiscale attention mechanism captures relationships between features at different scales. To address feature loss during transfer, the global information modeling module extracts global contextual features and suppresses background noise. In addition, the branch feature fusion module efficiently fuses valuable features from different levels to improve segmentation accuracy and confidence. To verify the effectiveness of CSFINet, we conducted extensive experiments on the underwater real-scene sonar image dataset. Specifically, our method achieved the mean intersection over union of 82.84$\%$ and the mean pixel accuracy of 89.37$\%$, outperforming several state-of-the-art methods, including convolutional neural networkss-based, Transformer-based, and Mamba-based models.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001432763400019  
Indexed Date:  
2025-03-06  
Institution Address:  
Xijing Univ, Sch Elect Informat, Xian 710123, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Sch Comp Sci, Xian 710129, Peoples R China  
Institution Address:  
Hohai Univ, Sch Earth Sci & Engn, Nanjing 210098, Peoples R China  
Institution Address:  
Air Force Engn Univ, Sch Informat & Nav, Xian 710082, Peoples R China  
Institution Address:  
Guangxi Acad Sci, Nanning 530007, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 401. GMRE-iUnet: Isomorphic Unet fusion model for PET and CT lung tumor

摘要: Lung tumor PET and CT image fusion is a key technology in clinical diagnosis. However, the existing fusion methods are difficult to obtain fused images with high contrast, prominent morphological features, and accurate spatial localization. In this paper, an isomorphic Unet fusion model (GMRE-iUnet) for lung tumor PET and CT images is proposed to address the above problems. The main idea of this network is as following: Firstly, this paper constructs an isomorphic Unet fusion network, which contains two independent multiscale dual encoders Unet, it can capture the features of the lesion region, spatial localization, and enrich the morphological information. Secondly, a Hybrid CNN-Transformer feature extraction module (HCTrans) is constructed to effectively integrate local lesion features and global contextual information. In addition, the residual axial attention feature compensation module (RAAFC) is embedded into the Unet to capture fine-grained information as compensation features, which makes the model focus on local connections in neighboring pixels. Thirdly, a hybrid attentional feature fusion module (HAFF) is designed for multiscale feature information fusion, it aggregates edge information and detail representations using local entropy and Gaussian filtering. Finally, the experiment results on the multimodal lung tumor medical image dataset show that the model in this paper can achieve excellent fusion performance compared with other eight fusion models. In CT mediastinal window images and PET images comparison experiment, AG, EI, QAB/F, SF, SD, and IE indexes are improved by 16.19%, 26%, 3.81%, 1.65%, 3.91% and 8.01%, respectively. GMRE-iUnet can highlight the information and morphological features of the lesion areas and provide practical help for the aided diagnosis of lung tumors.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
4  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:001096477700001  
PubMed ID:  
37826951  
Indexed Date:  
2023-11-26  
Institution Address:  
North Minzu Univ, Sch Comp Sci & Engn, Yinchuan 750021, Peoples R China  
Institution Address:  
Ningxia Med Univ, Sch Med Informat & Engn, Yinchuan 750004, Peoples R China  
Institution Address:  
Univ Leicester, Sch Comp & Math Sci, Leicester LE1 7RH, Leics, England  
Institution Address:  
North Minzu Univ, Key Lab Image & Graph Intelligent Processing, State Ethn Affairs Commiss, Yinchuan 750021, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 402. Optimal Deep Convolutional Neural Network for Vehicle Detection in

摘要: Object detection (OD) in remote sensing images (RSI) acts as a vital part in numerous civilian and military application areas, like urban plan-ning, geographic information system (GIS), and search and rescue functions. Vehicle recognition from RSIs remained a challenging process because of the difficulty of background data and the redundancy of recognition regions. The latest advancements in deep learning (DL) approaches permit the design of effectual OD approaches. This study develops an Artificial Ecosystem Optimizer with Deep Convolutional Neural Network for Vehicle Detec-tion (AEODCNN-VD) model on Remote Sensing Images. The proposed AEODCNN-VD model focuses on the identification of vehicles accurately and rapidly. To detect vehicles, the presented AEODCNN-VD model employs single shot detector (SSD) with Inception network as a baseline model. In addition, Multiway Feature Pyramid Network (MFPN) is used for handling objects of varying sizes in RSIs. The features from the Inception model are passed into the MFPN for multiway and multiscale feature fusion. Finally, the fused features are passed into bounding box and class prediction networks. For enhancing the detection efficiency of the AEODCNN-VD approach, AEO based hyperparameter optimizer is used, which is stimulated by the energy transfer strategies such as production, consumption, and decomposition in an ecosystem. The performance validation of the presented method on bench-mark datasets showed promising performance over recent DL models.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:000916172400017  
Indexed Date:  
2023-02-24  
Institution Address:  
Shaqra Univ, Coll Comp & Informat Technol, Dept Comp Sci, Shaqra, Saudi Arabia  
Institution Address:  
Umm Al Qura Univ, Coll Comp & Informat Syst, Dept Informat Syst, Mecca, Saudi Arabia  
Institution Address:  
Princess Nourah Bint Abdulrahman Univ, Coll Comp & Informat Sci, Dept Informat Syst, POB 84428, Riyadh 11671, Saudi Arabia  
Institution Address:  
Future Univ Egypt, Fac Engn & Technol, Dept Elect Engn, New Cairo 11845, Egypt  
Institution Address:  
Prince Sattam Bin Abdulaziz Univ, Dept Comp & Self Dev, Alkharj, Saudi Arabia  
Institution Address:  
Prince Sattam Bin Abdulaziz Univ, Coll Comp Engn & Sci, Dept Informat Syst, Alkharj, Saudi Arabia  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 403. MS-IAF: Multi-Scale Information Augmentation Framework for Aircraft

摘要: Aircrafts have been an important object of study in the field of multi-scale image object detection due to their important strategic role. However, the multi-scale detection of aircrafts and their key parts from remote sensing images can be a challenge, as images often present complex backgrounds and obscured conditions. Most of today's multi-scale datasets consist of independent objects and lack mixed annotations of aircrafts and their key parts. In this paper, we contribute a multi-scale aircraft dataset (AP-DATA) consisting of 7000 aircraft images that were taken in complex environments and obscured conditions. Our dataset includes mixed annotations of aircrafts and their key parts. We also present a multi-scale information augmentation framework (MS-IAF) to recognize multi-scale aircrafts and their key parts accurately. First, we propose a new deep convolutional module ResNeSt-D as the backbone, which stacks scattered attention in a multi-path manner and makes the receptive field more suitable for the object. Then, based on the combination of Faster R-CNN with ResNeSt-D, we propose a multi-scale feature fusion module called BFPCAR. BFPCAR overcomes the attention imbalance problem of the non-adjacent layers of the FPN module by reducing the loss of information between different layers and including more semantic features during information fusion. Based on AP-DATA, a dataset with three types of features, the average precision (AP) of MS-IAF reached 0.884, i.e., 2.67% higher than that of the original Faster R-CNN. The APs of these two modules were improved by 2.32% and 1.39%, respectively. The robustness of our proposed model was validated using the open sourced RSOD remote sensing image dataset, and the best accuracy was achieved.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
7  
eISSN:  
2072-4292  
入藏号:  
WOS:000839728800001  
Indexed Date:  
2022-08-23  
Institution Address:  
Northeastern Univ, Sensor & Big Data Lab, Qinhuangdao 066000, Hebei, Peoples R China  
Institution Address:  
Hebei Key Lab Micronano Precis Opt Sensing & Meas, Qinhuangdao 066000, Hebei, Peoples R China  
Institution Address:  
Army Engn Univ PLA, Shijiazhuang Sch, Shijiazhuang 050003, Hebei, Peoples R China  
Institution:  
Institution:  
Hebei Key Lab Micronano Precis Opt Sensing & Meas  
Institution:

### 404. Research on Wind Turbine Unbalance Fault Diagnosis Based on Wavelet

摘要: Rotor imbalance in wind turbines presents a serious problem. Particularly for offshore wind turbines, aerodynamic imbalance could have a severe impact because of the large size of the rotor. A diagnosis method based on a parallel convolutional neural network with multi-scale feature fusion is proposed to diagnose rotor imbalance. It consists of two feature extractors of different scales, which are combined in the fully connected layer. Firstly, a model of a 3MW wind turbine is built and the mass imbalance and aerodynamic imbalance are added to the simulation. The signal is collected and the effects of rotor imbalance on the nacelle vibration in wind turbines are investigated and described. Secondly, the nacelle vibration is selected as the target signal. Wavelet transform is performed on the collected signals, and the 2-dimensional time-frequency map is obtained as the object dataset for the classification. Thirdly, a convolutional neural network is used to classify rotor imbalances of different magnitudes, and different convolution kernels and activation functions are tested. Finally, a new data set is built in the highly fidelity simulation model, and the trained model is loaded for test and verification. The experiments show that the proposed diagnosis model based on the time-frequency map of nacelle vibrations and a convolutional neural network can identify rotor imbalance effectively, and the accuracy is greater than 98%. The results demonstrate the satisfactory performance of the proposed method. It can diagnose rotor imbalance effectively without additional sensors.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
2169-3536  
入藏号:  
WOS:001370660700036  
Indexed Date:  
2024-12-19  
Institution Address:  
CGN New Energy Holdings Co Ltd, Nanning 518031, Guangxi, Peoples R China  
Institution:  
CGN New Energy Holdings Co Ltd

### 405. DMAF-NET: Deep Multi-Scale Attention Fusion Network for Hyperspectral

摘要: In recent years, deep learning methods have achieved remarkable success in hyperspectral image classification (HSIC), and the utilization of convolutional neural networks (CNNs) has proven to be highly effective. However, there are still several critical issues that need to be addressed in the HSIC task, such as the lack of labeled training samples, which constrains the classification accuracy and generalization ability of CNNs. To address this problem, a deep multi-scale attention fusion network (DMAF-NET) is proposed in this paper. This network is based on multi-scale features and fully exploits the deep features of samples from multiple levels and different perspectives with an aim to enhance HSIC results using limited samples. The innovation of this article is mainly reflected in three aspects: Firstly, a novel baseline network for multi-scale feature extraction is designed with a pyramid structure and densely connected 3D octave convolutional network enabling the extraction of deep-level information from features at different granularities. Secondly, a multi-scale spatial-spectral attention module and a pyramidal multi-scale channel attention module are designed, respectively. This allows modeling of the comprehensive dependencies of coordinates and directions, local and global, in four dimensions. Finally, a multi-attention fusion module is designed to effectively combine feature mappings extracted from multiple branches. Extensive experiments on four popular datasets demonstrate that the proposed method can achieve high classification accuracy even with fewer labeled samples.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
1424-8220  
入藏号:  
WOS:001231640800001  
PubMed ID:  
38794007  
Indexed Date:  
2024-05-31  
Institution Address:  
North Univ China, Sch Instrument & Elect, State Key Lab Dynam Measurement Technol, Taiyuan 030051, Peoples R China  
Institution Address:  
Henan Coll Transportat, Dept Transportat Informat Engn, Zhengzhou 451460, Peoples R China  
Institution:  
Institution:  
Henan Coll Transportat

### 406. Wearable Sensor-Based Residual Multifeature Fusion Shrinkage Networks

摘要: Human activity recognition (HAR) based on wearable sensors has emerged as a low-cost key-enabling technology for applications such as human-computer interaction and healthcare. In wearable sensor-based HAR, deep learning is desired for extracting human active features. Due to the spatiotemporal dynamic of human activity, a special deep learning network for recognizing the temporal continuous activities of humans is required to improve the recognition accuracy for supporting advanced HAR applications. To this end, a residual multifeature fusion shrinkage network (RMFSN) is proposed. The RMFSN is an improved residual network which consists of a multi-branch framework, a channel attention shrinkage block (CASB), and a classifier network. The special multi-branch framework utilizes a 1D-CNN, a lightweight temporal attention mechanism, and a multi-scale feature extraction method to capture diverse activity features via multiple branches. The CASB is proposed to automatically select key features from the diverse features for each activity, and the classifier network outputs the final recognition results. Experimental results have shown that the accuracy of the proposed RMFSN for the public datasets UCI-HAR, WISDM, and OPPORTUNITY are 98.13%, 98.35%, and 93.89%, respectively. In comparison with existing advanced methods, the proposed RMFSN could achieve higher accuracy while requiring fewer model parameters.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001160398000001  
PubMed ID:  
38339474  
Indexed Date:  
2024-02-19  
Institution Address:  
Guangdong Polytech Normal Univ, Sch Elect & Informat, Guangzhou 510660, Peoples R China  
Institution:

### 407. Learning Higher Quality Rotation Invariance Features for Multioriented

摘要: Multioriented object detection, an important yet challenging task because of the bird's-eye-view perspective, complex background, and densely packed objects, is in the spotlight of detection in remote sensing images. Although existing methods have recently experienced substantial progress based on oriented head, they learn little about essential rotation invariance of the objects. In this article, a novel framework is proposed that can learn high-quality rotation invariance features of the multioriented objects by three measures. Given a remote sensing image, the multiscale semantic segmentation feature fusion module first merges the global semantic segmentation features predicted by the semantic segmentation branch and the multiscale features extracted by the backbone with FPN in order to distinguish complex background. Then, the discriminative features are used by rotation mainstream, whose structure is similar to cascade R-CNN and can extract higher quality rotation invariance features and predict more accurate location information by adaptively adjusting the distribution of the samples through progressive intersection over union thresholds. And in order to improve the performance of mainstream to predict more accurate oriented bounding box, the horizontal tributaries that can fully leverage the reciprocal relationship between the oriented detection and horizontal detection were added to the latter two stages. Extensive experiments on three public datasets for remote sensing images, i.e., Gaofen Airplane, HRSC2016, and DOTA demonstrate that without bells and whistles, the proposed method achieves superior performances compared with the existing state-of-the-art methods for multioriented detection. Moreover, our overall system achieves 59.264% mAP of airplane Detection in 2020 Gaofen challenge, ranking third in the final.  
在 BIOSIS Citation Index 中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
被引频次合计:  
10  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:000663535500016  
Indexed Date:  
2021-07-09  
Institution Address:  
Natl Univ Def Technol, State Key Lab Complex Electromagnet Environm Effe, Changsha 410073, Peoples R China  
Institution:

### 408. Attention V-Net: A Modified V-Net Architecture for Left Atrial

摘要: We propose a fully convolutional neural network based on the attention mechanism for 3D medical image segmentation tasks. It can adaptively learn to highlight the salient features of images that are useful for image segmentation tasks. Some prior methods enhance accuracy using multi-scale feature fusion or dilated convolution, which is basically artificial and lacks the flexibility of the model itself. Therefore, some works proposed the 2D attention gate module, but these works process 2D medical slice images, ignoring the correlation between 3D image sequences. In contrast, the 3D attention gate can comprehensively use the information of three dimensions of medical images. In this paper, we propose the Attention V-Net architecture, which uses the 3D attention gate module, and applied it to the left atrium segmentation framework based on semi-supervised learning. The proposed method is evaluated on the dataset of the 2018 left atrial challenge. The experimental results show that the Attention V-Net obtains improved performance under evaluation indicators, such as Dice, Jaccard, ASD (Average surface distance), and 95HD (Hausdorff distance). The result indicates that the model in this paper can effectively improve the accuracy of left atrial segmentation, therefore laying the foundation for subsequent work such as in atrial reconstruction. Meanwhile, our model is of great significance for assisting doctors in treating cardiovascular diseases.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
eISSN:  
2076-3417  
入藏号:  
WOS:000786248600001  
Indexed Date:  
2022-04-29  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Artificial Intelligence, Beijing 100876, Peoples R China  
Institution:

### 409. AF-EMS Detector: Improve the Multi-Scale Detection Performance of the

摘要: As a precursor step for computer vision algorithms, object detection plays an important role in various practical application scenarios. With the objects to be detected becoming more complex, the problem of multi-scale object detection has attracted more and more attention, especially in the field of remote sensing detection. Early convolutional neural network detection algorithms are mostly based on artificially preset anchor-boxes to divide different regions in the image, and then obtain the prior position of the target. However, the anchor box is difficult to set reasonably and will cause a large amount of computational redundancy, which affects the generality of the detection model obtained under fixed parameters. In the past two years, anchor-free detection algorithm has achieved remarkable development in the field of detection on natural image. However, there is no sufficient research on how to deal with multi-scale detection more effectively in anchor-free framework and use these detectors on remote sensing images. In this paper, we propose a specific-attention Feature Pyramid Network (FPN) module, which is able to generate a feature pyramid, basing on the characteristics of objects with various sizes. In addition, this pyramid suits multi-scale object detection better. Besides, a scale-aware detection head is proposed which contains a multi-receptive feature fusion module and a size-based feature compensation module. The new anchor-free detector can obtain a more effective multi-scale feature expression. Experiments on challenging datasets show that our approach performs favorably against other methods in terms of the multi-scale object detection performance.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
被引频次合计:  
11  
eISSN:  
2072-4292  
入藏号:  
WOS:000611558400001  
Indexed Date:  
2021-03-10  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100190, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Key Lab Network Informat Syst Technol NIST, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100190, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 100190, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 410. An Improved Multi-Scale Feature Fusion for Skin Lesion Segmentation

摘要: Accurate segmentation of skin lesions is still a challenging task for automatic diagnostic systems because of the significant shape variations and blurred boundaries of the lesions. This paper proposes a multi-scale convolutional neural network, REDAUNet, based on UNet3+ to enhance network performance for practical applications in skin segmentation. First, the network employs a new encoder module composed of four feature extraction layers through two cross-residual (CR) units. This configuration allows the module to extract deep semantic information while avoiding gradient vanishing problems. Subsequently, a lightweight and efficient channel attention (ECA) module is introduced during the encoder's feature extraction stage. The attention module assigns suitable weights to channels through attention learning and effectively captures inter-channel interaction information. Finally, the densely connected atrous spatial pyramid pooling module (DenseASPP) module is inserted between the encoder and decoder paths. This module integrates dense connections and ASPP, as well as multi-scale information fusion, to recognize lesions of varying sizes. The experimental studies in this paper were constructed on two public skin lesion datasets, namely, ISIC-2018 and ISIC-2017. The experimental results show that our model is more accurate in segmenting lesions of different shapes and achieves state-of-the-art performance in segmentation. In comparison to UNet3+, the proposed REDAUNet model shows improvements of 2.01%, 4.33%, and 2.68% in Dice, Spec, and mIoU metrics, respectively. These results suggest that REDAUNet is well-suited for skin lesion segmentation and can be effectively employed in computer-aided systems.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
8  
被引频次合计:  
8  
eISSN:  
2076-3417  
入藏号:  
WOS:001034938700001  
Indexed Date:  
2023-08-10  
Institution Address:  
Univ Sci & Technol Liaoning, Sch Comp Sci & Software Engn, Anshan 114051, Peoples R China  
Institution:

### 411. Geometric prior guided hybrid deep neural network for facial beauty

摘要: Facial beauty analysis is an important topic in human society. It may be used as a guidance for face beautification applications such as cosmetic surgery. Deep neural networks (DNNs) have recently been adopted for facial beauty analysis and have achieved remarkable performance. However, most existing DNN-based models regard facial beauty analysis as a normal classification task. They ignore important prior knowledge in traditional machine learning models which illustrate the significant contribution of the geometric features in facial beauty analysis. To be specific, landmarks of the whole face and facial organs are introduced to extract geometric features to make the decision. Inspired by this, we introduce a novel dual-branch network for facial beauty analysis: one branch takes the Swin Transformer as the backbone to model the full face and global patterns, and another branch focuses on the masked facial organs with the residual network to model the local patterns of certain facial parts. Additionally, the designed multi-scale feature fusion module can further facilitate our network to learn complementary semantic information between the two branches. In model optimisation, we propose a hybrid loss function, where especially geometric regulation is introduced by regressing the facial landmarks and it can force the extracted features to convey facial geometric features. Experiments performed on the SCUT-FBP5500 dataset and the SCUT-FBP dataset demonstrate that our model outperforms the state-of-the-art convolutional neural networks models, which proves the effectiveness of the proposed geometric regularisation and dual-branch structure with the hybrid network. To the best of our knowledge, this is the first study to introduce a Vision Transformer into the facial beauty analysis task.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
2468-6557  
eISSN:  
2468-2322  
入藏号:  
WOS:000949583300001  
Indexed Date:  
2023-04-05  
Institution Address:  
Guizhou Univ, Sch Comp Sci & Technol, Guiyang, Peoples R China  
Institution Address:  
Moutai Inst, Dept Automat, Renhuai, Guizhou, Peoples R China  
Institution Address:  
Harbin Inst Technol, Sch Comp Sci & Technol, Shenzhen, Peoples R China  
Institution Address:  
Dalian Minzu Univ, Informat & Commun Engn Dept, Dalian, Peoples R China  
Institution Address:  
Chinese Univ Hong Kong, Sch Data Sci, Shenzhen, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 412. MSF-CSPNet: A Specially Designed Backbone Network for Faster R-CNN

摘要: Although Faster R-CNN has undergone a lot of improvements, it still exists a significant gap in the performance between the detection of small and large objects, mainly because the low-level network lacks semantic information and small objects are only involved in a few images. To mitigate the above issues, we propose an object detection model based on Multi-Scale Feature fusion Cross Stage Partial Network (MSF-CSPNet) in this paper. The proposed MSF-CSPNet focuses on the fusion of concrete features and abstract features from multi-scale feature by learning shallow features at the shallow level and deep features at the deep level. Meanwhile, the data augmentation is performed by using random horizontal flip. On the basis, the improved Faster-RCNN model with Automatic Mixed Precision, Group Batch Sampler and MSF-CSPNet was formed. The proposed algorithm is valuated on the Microsoft Common Objects in Context (MS COCO) 2017 and obtained leading performance with 5.4% improvement in APcoco, 5.9% improvement in AP50, 6.9% improvement in AP75, 5.8% improvement in APS, 6.1% improvement in APM, 5.8% improvement in APL compare to Faster R-CNN based on ResNet-50 with Feature Pyramid Network (FPN) backbone, and also outperformed previous reports on state-of-art Faster R-CNN series using other backbone networks, especially for small object detection. This research shows that the combination of a backbone with stronger learning ability and FPN is helpful to detect the expression of objects. Faster R-CNN based on MSF-CSPNet has high efficiency and better balance between accuracy and speed.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
2169-3536  
入藏号:  
WOS:001204906900001  
Indexed Date:  
2024-04-25  
Institution Address:  
Henan Agr Univ, Coll Informat & Management Sci, Zhengzhou 450046, Peoples R China  
Institution:

### 413. MsFNet: Multi-Scale Fusion Network Based on Dynamic Spectral Features

摘要: With the development of satellite technology, the importance of multi-temporal remote sensing (RS) image change detection (CD) in urban planning, environmental monitoring, and other fields is increasingly prominent. Deep learning techniques enable a profound exploration of the intrinsic features within hyperspectral (HS) data, leading to substantial enhancements in CD accuracy while addressing several challenges posed by traditional methodologies. However, existing convolutional neural network (CNN)-based CD approaches frequently encounter issues during the feature extraction process, such as the loss of detailed information due to downsampling, which hampers a model's ability to accurately capture complex spectral features. Additionally, these methods often neglect the integration of multi-scale information, resulting in suboptimal local feature extraction and, consequently, diminished model performance. To address these limitations, we propose a multi-scale fusion network (MsFNet) which leverages dynamic spectral features for effective multi-temporal HS-CD. Our approach incorporates a dynamic convolution module with spectral attention, which adaptively modulates the receptive field size according to the spectral characteristics of different bands. This flexibility enhances the model's capacity to focus on critical bands, thereby improving its ability to identify and differentiate changes across spectral dimensions. Furthermore, we develop a multi-scale feature fusion module which extracts and integrates features from deep feature maps, enriching local information and augmenting the model's sensitivity to local variations. Experimental evaluations conducted on three real-world HS-CD datasets demonstrate that the proposed MsFNet significantly outperforms contemporary advanced CD methods in terms of both efficacy and performance.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001305253700001  
Indexed Date:  
2024-09-09  
Institution Address:  
Liaoning Normal Univ, Sch Geog, Dalian 116029, Peoples R China  
Institution Address:  
Liaoning Normal Univ, Sch Comp Sci & Artificial Intelligence, Dalian 116029, Peoples R China  
Institution:  
Institution:

### 414. Source camera identification based on an adaptive dual-branch fusion

摘要: Although deep learning algorithms have addressed the issue of identifying the source camera to a certain extent, developing a straightforward and effective network remains a challenging task. At present, most of the excellent network schemes in source camera identification are deep networks, which heavily rely on the strong feature extraction ability of deep networks. Although deepening network layers has achieved certain results, training a deep convolutional neural network model requires a large dataset, sophisticated hardware and lengthy training time, and there is a waste of resources. To solve the problem of redundant structure and resource waste of deep convolutional neural networks, this paper proposes the SE-BRB module, which we call a new network module based on the residual module and SE module. Based on this, an adaptive dual-branch fusion network (ADF-Net) with a simplified structure is designed to identify the source of digital images. Specifically, the bottleneck residual module can achieve direct backward transfer of shallow features to avoid images being over-compressed and is suitable for capturing weak source features in images; Additionally, the introduction of a channel attention mechanism can increase the weight of effective feature channels in the network and improve network performance. Finally, multiscale camera feature fusion is realized through a dual-branch network structure to further improve the network performance. The accuracy of the model proposed in this paper is 99.33% and 98.78% on the Dresden dataset and the self-built complex dataset, respectively, and the classification accuracy is far ahead of the existing source camera identification methods.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001032628900001  
Indexed Date:  
2023-08-02  
Institution Address:  
Wuhan Univ, Sch Elect Informat, Wuhan 430000, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Cyber Sci & Engn, Wuhan 430072, Peoples R China  
Institution Address:  
Hubei Business Coll, Sch Artificial Intelligence, Wuhan 430079, Peoples R China  
Institution Address:  
Nanchang Inst Sci & Technol, Coll Artificial Intelligence, Nanchang 330108, Peoples R China  
Institution:  
Institution:  
Institution:  
Hubei Business Coll  
Institution:  
Nanchang Inst Sci & Technol

### 415. Hyperspectral image change detection based on an improved multi-scale

摘要: Change detection (CD) has a significant application in the remote sensing field. Because of the popularity of hyperspectral image (HSI) and the application of deep learning methods, hyperspectral image change detection (HSI-CD) techniques have been greatly developed. Among them, convolutional neural network (CNN) has garnered the greatest interest in HSI-CD due to their superior feature learning capabilities. However, current CNN-based algorithms have trouble capturing spectral similarity and long-range dependency owing to their intrinsic structural restrictions. Recently, transformers have been shown to extract global dependency from nature images in an extremely efficient way. But it has some difficulties in handling high-dimensional data, such as HSI. To address these issues, we propose an improved multi-scale and spectral-wise transformer (MS-SWT). The proposed MS-SWT is capable of capturing spectral similarity and long-range dependence between bands to enhance the efficiency of the HSI-CD task. Furthermore, to maximize the utilization of spatial information, we present a multi-scale feature fusion module (MFFM) to extract and fuse different dimensions of spatial features. More importantly, a locality self-attention (LSA) module is employed to alleviate the problem of smoothing the distribution of attention scores due to the large number of spectral embeddings. Moreover, we design a channel self-supervised loss function that can capture intrinsic information from the spectral channels to further strengthen the robustness of model training when the training samples are scarce. Lastly, comprehensive experiments present the high performance of our MS-SWT on four bitemporal HSI datasets and demonstrate the superiority of MS-SWT over state-of-the-art approaches.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:001176348300001  
Indexed Date:  
2024-03-22  
Institution Address:  
Fuzhou Univ, Coll Comp & Data Sci, Fuzhou, Peoples R China  
Institution Address:  
Fuzhou Univ, Coll Math & Stat, Fuzhou, Peoples R China  
Institution:  
Institution:

### 416. A swin-transformer-based network with inductive bias ability for medical

摘要: Accurately segmenting organs, or diseased regions of varying sizes, is a challenge in medical image segmentation tasks with limited datasets. Although Transformer-based methods have self-attention mechanisms, excellent global modelling abilities and can effectively focus on crucial areas in medical images, they still face the intractable issues of computational complexity and excessive reliance on data. The Swin-Transformer has partly addressed the problem of computational complexity, however the method still requires large amounts of training data due to its lack of inductive bias capabilities. When applied to medical image segmentation tasks with small datasets, this leads to suboptimal performances. In contrast, the CNN-based methods can compensate for this limitation. To address this issue further, this paper proposes Swin-IBNet, which combines the Swin-Transformer with a CNN in a novel manner to imbue it with inductive bias capabilities, reducing its reliance on data. During the encoding process of Swin-IBNet, two novel and crucial modules, the feature fusion block (FFB) and the multiscale feature aggregation block (MSFA), are designed. The FFB is responsible for propagating the inductive bias capability to the Swin-Transformer encoder. Different from the previous use of multiscale features, MSFA efficiently leverages multiscale information from different layers through self-learning. This paper not only attempts to analyse the interpretability of the proposed Swin-IBNet but also performs more verifications on the public Synapse, ISIC 2018 and ACDC datasets. The experimental results show that Swin-IBNet is superior to the baseline method, Swin-Unet, and several state-of-the-art methods. Especially on the Synapse dataset, the DSC of Swin-IBNet surpasses that of Swin-Unet by 3.45%.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:001372904500004  
Indexed Date:  
2024-12-16  
Institution Address:  
Jilin Univ, Coll Comp Sci & Technol, Key Lab Symbol Computat & Knowledge Engn, Minist Educ, Changchun 130012, Jilin, Peoples R China  
Institution:

### 417. Small-Target Detection Based on an Attention Mechanism for

摘要: Small-target detection suffers from the problems of low average precision and difficulties detecting targets from airport-surface surveillance videos. To address this challenge, this study proposes a small-target detection model based on an attention mechanism. First, a standard airport small-target dataset was established, where the absolute scale of each marked target meets the definition of a small target. Second, using the Mask Scoring R-CNN model as a baseline, an attention module was added to the feature extraction network to enhance its feature representation and improve the accuracy of its small-target detection. A multiscale feature pyramid fusion module was used to fuse more detailed shallow information according to the feature differences of diverse small targets. Finally, a more effective detection branch structure is proposed to improve detection accuracy. Experimental results verify the effectiveness of the proposed method in detecting small targets. Compared to the Mask R-CNN and Mask Scoring R-CNN models, the detection accuracy of the proposed method in two-pixel intervals with the lowest rate of small targets increased by 10%, 3.04% and 16%, 15.15%, respectively. The proposed method proved to have a higher accuracy and be more effective at small-target detection.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
4  
eISSN:  
2076-3417  
入藏号:  
WOS:000986897600001  
Indexed Date:  
2023-05-31  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Civil Aviat, Nanjing 210016, Peoples R China  
Institution Address:  
China Acad Civil Aviat Sci & Technol, Beijing 100028, Peoples R China  
Institution:  
Institution:  
China Acad Civil Aviat Sci & Technol

### 418. D2-CovidNet: A Deep Learning Model for COVID-19 Detection in Chest X-Ray

摘要: Since the outbreak of Coronavirus disease 2019 (COVID-19), it has been spreading rapidly worldwide and has not yet been effectively controlled. Many researchers are studying novel Coronavirus pneumonia from chest X-ray images. In order to improve the detection accuracy, two modules sensitive to feature information, dual-path multiscale feature fusion module and dense depthwise separable convolution module, are proposed. Based on these two modules, a lightweight convolutional neural network model, D2-CovidNet, is designed to assist experts in diagnosing COVID-19 by identifying chest X-ray images. D2-CovidNet is tested on two public data sets, and its classification accuracy, precision, sensitivity, specificity, and F1-score are 94.56%, 95.14%, 94.02%, 96.61%, and 95.30%, respectively. Specifically, the precision, sensitivity, and specificity of the network for COVID-19 are 98.97%, 94.12%, and 99.84%, respectively. D2-CovidNet has fewer computation number and parameter number. Compared with other methods, D2-CovidNet can help diagnose COVID-19 more quickly and accurately.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1687-5265  
eISSN:  
1687-5273  
入藏号:  
WOS:000783255200005  
PubMed ID:  
34925500  
Indexed Date:  
2022-05-05  
Institution Address:  
Changsha Univ Sci & Technol, Sch Comp & Commun Engn, Changsha 410114, Peoples R China  
Institution Address:  
Hunan Childrens Hosp, Changsha 410000, Peoples R China  
Institution:  
Institution:  
Hunan Childrens Hosp

### 419. B-YOLOX-S: A Lightweight Method for Underwater Object Detection Based on

摘要: With the increasing maturity of underwater agents-related technologies, underwater object recognition algorithms based on underwater robots have become a current hotspot for academic and applied research. However, the existing underwater imaging conditions are poor, the images are blurry, and the underwater robot visual jitter and other factors lead to lower recognition precision and inaccurate positioning in underwater target detection. A YOLOX-based underwater object detection model, B-YOLOX-S, is proposed to detect marine organisms such as echinus, holothurians, starfish, and scallops. First, Poisson fusion is used for data amplification at the input to balance the number of detected targets. Then, wavelet transform is used to perform Style Transfer on the enhanced images to achieve image restoration. The clarity of the images and detection targets is further increased and the generalization of the model is enhanced. Second, a combination of BIFPN-S and FPN is proposed to fuse the effective feature layer obtained by the Backbone layer to enhance the detection precision and accelerate model detection. Finally, the localization loss function of the prediction layer in the network is replaced by EIoU\_Loss to heighten the localization precision in detection. Experimental results comparing the B-YOLOX-S algorithm model with mainstream algorithms such as FasterRCNN, YOLOV3, YOLOV4, YOLOV5, and YOLOX on the URPC2020 dataset show that the detection precision and detection speed of the algorithm model have obvious advantages over other algorithm networks. The average detection accuracy mAP value is 82.69%, which is 5.05% higher than the benchmark model (YOLOX-s), and the recall rate is 8.03% higher. Thus, the validity of the algorithmic model proposed in this paper is demonstrated.  
在 BIOSIS Citation Index 中的被引频次:  
3  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
被引频次合计:  
10  
eISSN:  
2077-1312  
入藏号:  
WOS:000895111800001  
Indexed Date:  
2022-12-22  
Institution Address:  
Taiyuan Univ Technol, Coll Software, Taiyuan 030024, Peoples R China  
Institution:

### 420. Swin-MRDB: Pan-Sharpening Model Based on the Swin Transformer and

摘要: Pan-sharpening aims to create high-resolution spectrum images by fusing low-resolution hyperspectral (HS) images with high-resolution panchromatic (PAN) images. Inspired by the Swin transformer used in image classification tasks, this research constructs a three-stream pan-sharpening network based on the Swin transformer and a multi-scale feature extraction module. Unlike the traditional convolutional neural network (CNN) pan-sharpening model, we use the Swin transformer to establish global connections with the image and combine it with a multi-scale feature extraction module to extract local features of different sizes. The model combines the advantages of the Swin transformer and CNN, enabling fused images to maintain good local detail and global linkage by mitigating distortion in hyperspectral images. In order to verify the effectiveness of the method, this paper evaluates fused images with subjective visual and quantitative indicators. Experimental results show that the method proposed in this paper can better preserve the spatial and spectral information of images compared to the classical and latest models.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001046201100001  
Indexed Date:  
2023-08-25  
Institution Address:  
Nanjing Forestry Univ, Sch Mech & Elect Engn, Nanjing 210037, Peoples R China  
Institution:

### 421. Vision Detection Method for Picking Robots Based on Improved Faster R -

摘要: To address the issue of poor detection and positioning capabilities of fruit picking robots in scenes with densely distributed targets and fruits occluding each other,a method to improve the fruit detection and positioning of Faster R - CNN was proposed by introducing an efficient channel attention mechanism(ECA) and a multiscale feature fusion pyramid(FPN). Firstly, the commonly used VGG16 network was replaced with a ResNet50 residual network with strong expression capability and eliminate network degradation problem, thus extracting more abstract and rich semantic information to enhance the model's detection ability for multiscale and small targets. Secondly, the ECA module was introduced to enable the feature extraction network to focus on local and efficient information in the feature map, reduce the interference of invalid targets,and improve the model's detection accuracy. Finally,a branch and leaf grafting data augmentation method was used to improve the apple dataset and solve the problem of insufficient image data. Based on the constructed dataset,genetic algorithms were used to optimize K-means + + clustering and generate adaptive anchor boxes. Experimental results showed that the improved model had average precision of 96.16% for graspable apples and 86.95% for non-graspable apples,and the mean average precision was 92.79%,which was 15.68 percentages higher than that of the traditional Faster R - CNN. The positioning accuracy for graspable and non-directly graspable apples were 97.14% and 88.93%,respectively,which were 12.53 percentages and 40.49 percentages higher than that of traditional Faster R - CNN. The weight was reduced by 38.20%. The computation time was reduced by 40.7%. The improved model was more suitable for application in fruit-picking robot visual systems.  
摘要:  
针对采摘机器人对场景中目标分布密集、果实相互遮挡的检测及定位能力不理想问题,提出一种引入高效通道注意力机制(ECA)和多尺度融合特征金字塔(FPN)改进Faster R - CNN果实检测及定位方法。首先,利用表达能力较强的融合FPN的残差网络ResNet50替换原VGG16网络,消除了网络退化问题,进而提取更加抽象和丰富的语义信息,提升模型对多尺度和小目标的检测能力;其次,引入注意力机制ECA模块,使特征提取网络聚焦特征图像的局部高效信息,减少无效目标的干扰,提升模型检测精度;最后,采用一种枝叶插图数据增强方法改进苹果数据集,解决图像数据不足问题。基于构建的数据集,使用遗传算法优化K-means + +聚类生成自适应锚框,提高模型定位准确性。试验结果表明,改进模型对可抓取和不可直接抓取苹果的精度均值分别为96.16%和86.95%,平均精度均值为92.79%,较传统Faster R - CNN提升15.68个百分点;对可抓取和不可直接抓取的苹果定位精度分别为97.14%和88.93%,较传统Faster R - CNN分别提高12.53个百分点和40.49个百分点;内存占用量减少38.20%,每帧平均计算时间缩短40.7%,改进后的模型参数量小且实时性好,能够更好地应用于果实采摘机器人视觉系统。  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1000-1298  
入藏号:  
CSCD:7647452  
Indexed Date:  
2024-04-12  
Institution Address:  
兰州理工大学机电工程学院, 兰州, 甘肃 730050, 中国  
Institution Address:  
School of Mechanical and Electrical Engineering,Lanzhou University of Technology, Lanzhou, Gansu 730050, China  
Institution:  
兰州理工大学机电工程学院  
Institution:  
School of Mechanical and Electrical Engineering,Lanzhou University of Technology

### 422. Pulmonary Nodule Detection Based on Multiscale Feature Fusion.

摘要: As cancer with the highest morbidity and mortality in the world, lung cancer is characterized by pulmonary nodules in the early stage. The detection of pulmonary nodules is an important method for the early detection of lung cancer, which can greatly improve the survival rate of lung cancer patients. However, the accuracy of conventional detection methods for lung nodules is low. With the development of medical imaging technology, deep learning plays an increasingly important role in medical image detection, and pulmonary nodules can be accurately detected by CT images. Based on the above, a pulmonary nodule detection method based on deep learning is proposed. In the candidate nodule detection stage, the multiscale features and Faster R-CNN, a general-purpose detection framework based on deep learning, were combined together to improve the detection of small-sized lung nodules. In the false-positive nodule filtration stage, a 3D convolutional neural network based on multiscale fusion is designed to reduce false-positive nodules. The experiment results show that the candidate nodule detection model based on Faster R-CNN integrating multiscale features has achieved a sensitivity of 98.6%, 10% higher than that of the other single-scale model, the proposed method achieved a sensitivity of 90.5% at the level of 4 false-positive nodules per scan, and the CPM score reached 0.829. The results are higher than methods in other works of literature. It can be seen that the detection method of pulmonary nodules based on multiscale fusion has a higher detection rate for small nodules and improves the classification performance of true and false-positive pulmonary nodules. This will help doctors when making a lung cancer diagnosis.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
eISSN:  
1748-6718  
入藏号:  
MEDLINE:36590762  
PubMed ID:  
36590762  
Indexed Date:  
2023-01-03  
Institution Address:  
College of Medicine & Biological Information Engineering, Northeastern University, Shenyang 110169, China.  
Institution Address:  
School of Computer Science & Engineering, Northeastern University, Shenyang 110169, China.  
Institution Address:  
Key Laboratory of Big Data Management and Analytics, Liaoning Province, Northeastern University, Shenyang 110169, China.  
Institution:  
Institution:  
Institution:

### 423. A lightweight improved YOLOv5s model and its deployment for detecting

摘要: Precise detection and low-cost deployment are the technological basis of intelligent fruit picking. This study proposes a lightweight improved YOLOv5s model to detect pitaya fruits in daytime and nighttime lightsupplement environments, and make it successfully deploy in an Android device. This model first uses the module of shufflenetv2 to reconstruct the YOLOv5s backbone network. Then, the study proposes a ConcentratedComprehensive Convolution Receptive Field Enhancement (C3RFE) module to improve the detection precision of pitaya fruits. Furthermore, a Bidirectional Feature Pyramid Network (BiFPN) feature fusion method is used to enhance the multi-scale feature fusion. Moreover, three optimized Squeeze-and-Excitation (SE) attention modules are added to make full use of the image feature information. Finally, a dynamic label allocation strategy simple Optimal Transport Assignment (simOTA) is used to optimize the YOLOv5s model original label allocation strategy. The experimental results show that the improved model achieves an average precision rate of 97.80 %, with frames per second (FPS) of 139 FPS in a GPU run environment. The model size is only 2.5 MB. Compared to the state-of-the-art SSD, Faster RCNN, YOLOv4, YOLOv4 tiny, YOLOv5s, YOLOv5Lite-s, YOLOXs, YOLOv7, YOLOv7-tiny, YOLOv8n and YOLOv8s, the improved YOLOv5s achieve preferred comprehensive performance in average precision rate, FPS and model size. When deploying this model on the Realme GT Android mobile phone by developing an application based on an NCNN framework, such an application accomplishes real-time pitaya fruit detection with an FPS exceeding 30 FPS. This study can provide technological support for a precise and effective pitaya fruit intelligent picking.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
46  
在 BIOSIS Citation Index 中的被引频次:  
21  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
47  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:001225336200001  
Indexed Date:  
2024-05-23  
Institution Address:  
Guangxi Univ, Sch Mech Engn, 100 East Daxue Rd, Nanning 530004, Peoples R China  
Institution Address:  
South China Univ Technol, Sch Automat Sci & Engn, 381 Wushan Rd, Guangzhou 510641, Peoples R China  
Institution Address:  
Guangzhou Coll Commerce, Sch Modern Informat Ind, Guangzhou 510000, Peoples R China  
Institution Address:  
Guangxi Univ, Sch Mech Engn, Guangxi Key Lab Mfg Syst & Adv Mfg Technol, Nanning 530004, Peoples R China  
Institution Address:  
South China Agr Univ, Coll Engn, 483 Wushan Rd, Guangzhou 510642, Guangdong, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 424. Ultrasound spine image segmentation using multi-scale feature fusion

摘要: Scoliosis is a 3D spinal deformation where the spine takes a lateral curvature, forming an angle in the coronal plane. Diagnosis of scoliosis requires periodic detection, and frequent exposure to radiative imaging may cause cancer. A safer and more economical alternative imaging, i.e., 3D ultrasound imaging modality, is being explored. However, unlike other radiative modalities, an ultrasound image is noisy, which often suppresses the image's useful information. Through this research, a novel hybridized CNN architecture, multi-scale feature fusion Skip-Inception U-Net (SIU-Net), is proposed for a fully automatic bony feature detection, which can be further used to assess the severity of scoliosis safely and automatically. The proposed architecture, SIU-Net, incorporates two novel features into the basic U-Net architecture: (a) an improvised Inception block and (b) newly designed decoder-side dense skip pathways. The proposed model is tested on 109 spine ultrasound image datasets. The architecture is evaluated using the popular (i) Jaccard Index (ii) Dice Coefficient and (iii) Euclidean distance, and compared with (a) the basic U-net segmentation model, (b) a more evolved UNet++ model, and (c) a newly developed MultiResUNet model. The results show that SIU-Net gives the clearest segmentation output, especially in the important regions of interest such as thoracic and lumbar bony features. The method also gives the highest average Jaccard score of 0.781 and Dice score of 0.883 and the lowest histogram Euclidean distance of 0.011 than the other three models. SIU-Net looks promising to meet the objectives of a fully automatic scoliosis detection system.(c) 2022 Nalecz Institute of Biocybernetics and Biomedical Engineering of the Polish Academy of Sciences. Published by Elsevier B.V. All rights reserved.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
18  
在 BIOSIS Citation Index 中的被引频次:  
2  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
19  
ISSN:  
0208-5216  
入藏号:  
WOS:000784299100004  
Indexed Date:  
2022-04-28  
Institution Address:  
Univ Technol Sydney, Sch Biomed Engn, Sydney, NSW, Australia  
Institution Address:  
Harbin Engn Univ, Coll Informat & Commun Engn, Harbin, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Dept Elect & Informat Engn, Hong Kong, Peoples R China  
Institution Address:  
Hong Kong Polytech Univ, Dept Biomed Engn, Hong Kong, Peoples R China  
Institution Address:  
Univ Technol, Bldg 11,15 Broadway, Ultimo, NSW 2007, Australia  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 425. ConvMedSegNet: A multi-receptive field depthwise convolutional neural

摘要: In order to achieve highly precise medical image segmentation, this paper presents ConvMedSegNet, a novel convolutional neural network designed with a U-shaped architecture that seamlessly integrates two crucial modules: the multi-receptive field depthwise convolution module (MRDC) and the guided fusion module (GF). The MRDC module's primary function is to capture texture information of varying sizes through multi-scale convolutional layers. This information is subsequently utilized to enhance the correlation of global feature data by expanding the network's width. This strategy adeptly preserves the inherent inductive biases of convolution while concurrently amplifying the network's ability to establish dependencies on global information. Conversely, the GF module assumes responsibility for implementing multi-scale feature fusion by connecting the encoder and decoder components. It facilitates the transfer of information between features that are separated over substantial distances through guided fusion, effectively minimizing the loss of critical data. In experiments conducted on public medical image datasets such as BUSI and ISIC2018, ConvMedSegNet outperforms several advanced competing methods, yielding superior results. Additionally, the code can be accessed at https://github.com/csust-yixin/ConvMedSegNet.  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
1879-0534  
入藏号:  
MEDLINE:38759586  
PubMed ID:  
38759586  
Indexed Date:  
2024-05-19  
Institution Address:  
School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha, 410114, China.  
Institution Address:  
Huaihua University, Huaihua, 418000, China.  
Institution Address:  
Changkuangao Beijing Technology Co., Ltd, Beijing 101100, China.  
Institution Address:  
Wenzhou University Library, Wenzhou, 325035, China. Electronic address: zhouzf@wzu.edu.cn.  
Institution:  
Institution:  
Institution:  
Institution:

### 426. A novel diagnostic framework based on vibration image encoding and

摘要: Intelligent fault diagnosis employing CNN-based techniques has demonstrated promising results in rotating machinery maintenance and management. However, most approaches that rely on time-series vibration signals fail to incorporate the physical knowledge of faults into feature learning, and are subject to obtaining the correlation between input and output characteristics. Furthermore, it is challenging for the neural network to improve feature extraction due to the rigid convolution kernel and the limited receptive field. To address these issues, an intelligent fault diagnosis framework is developed based on vibration image encoding and multi-scale neural networks. Firstly, a novel image encoding rule is designed to convert time-series signals into vibration images, which incorporate the impact characteristics of faults into the model inputs to enhance fault information. Secondly, a multi-scale feature fusion network with dilatated and irregular sampling convolution block is constructed for feature extraction and classification. In addition, a new regularization term is designed to avoid training overfitting. Finally, Grad-CAM is introduced to visualize the region of interest of the model. The proposed method is verified on the datasets with different fault types from our laboratory and DIRG bearing dataset with different fault severities from Politecnico di Torino. The experimental results demonstrate that our diagnostic framework exhibits good generalization and noise robustness, and can achieve high-precision fault diagnosis.  
Web of Science 核心合集中的 "被引频次":  
7  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
7  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001232482800002  
Indexed Date:  
2024-05-31  
Institution Address:  
Yanshan Univ, Qinhuangdao, Peoples R China  
Institution:

### 427. Joint Rain Streaks & Haze Removal Network for Object Detection

摘要: In the realm of low-level vision tasks, such as image deraining and dehazing, restoring images distorted by adverse weather conditions remains a significant challenge. The emergence of abundant computational resources has driven the dominance of deep Convolutional Neural Networks (CNNs), supplanting traditional methods reliant on prior knowledge. However, the evolution of CNN architectures has tended towards increasing complexity, utilizing intricate structures to enhance performance, often at the expense of computational efficiency. In response, we propose the Selective Kernel Dense Residual M-shaped Network (SKDRMNet), a flexible solution adept at balancing computational efficiency with network accuracy. A key innovation is the incorporation of an Mshaped hierarchical structure, derived from the U-Net framework as M-Network (M-Net), within which the Selective Kernel Dense Residual Module (SDRM) is introduced to reinforce multi-scale semantic feature maps. Our methodology employs two sampling techniques-bilinear and pixel unshuffled and utilizes a multi-scale feature fusion approach to distil more robust spatial feature map information. During the reconstruction phase, feature maps of varying resolutions are seamlessly integrated, and the extracted features are effectively merged using the Selective Kernel Fusion Module (SKFM). Empirical results demonstrate the comprehensive superiority of SKDRMNet across both synthetic and real rain and haze datasets.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001268414500001  
Indexed Date:  
2024-07-22  
Institution Address:  
Natl Inst Technol, Dept Elect & Commun Engn, Warangal 506004, India  
Institution Address:  
Natl Inst Technol, Dept Comp Sci & Engn, Warangal 506004, India  
Institution Address:  
Indian Inst Informat Technol Design & Mfg, Dept Elect & Commun Engn, Kurnool 518008, India  
Institution:  
Institution:  
Institution:  
Indian Inst Informat Technol Design & Mfg

### 428. Image presentation and effective classification of odor intensity levels

摘要: This paper proposes a novel data processing scheme for electronic noses that combines the gramian angular field (GAF) and convolutional neural network (CNN) to achieve high performance in classifying five levels of odor intensity. Specifically, a multi-channel e-nose was developed to detect various gases, including hydrogen sulfide, ammonia, sulfur dioxide, trimethylamine, and alkane gases, among others, in complex odor components. The sensor array was optimized through Spearman correlation analysis of the sensor signals and artificial olfactory odor intensity levels. Moreover, the one-dimensional temporal sensor data was converted into two-dimensional color images using the GAF (GASF/GADF) algorithm. This approach enables a more detailed presentation of deep features while retaining the time-domain dependence of the signals. To enhance the performance of classification, a multi-scale feature fusion network (MFFNet) was designed. Notably, GASF-converted images are more effective in characterizing sensor data for different odor intensity levels than GADF-converted images. Compared to classical CNN classification models such as AlexNet, GoogLeNet, and ResNet18, MFFNet achieved the highest accuracy and macro average F1-score on the test set, which were 93.75% and 93.34%, respectively. The results demonstrate the efficient classification of odor intensity levels by combining multi-channel e-nose technology, GASF, and CNN.  
Web of Science 核心合集中的 "被引频次":  
19  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
21  
eISSN:  
0925-4005  
入藏号:  
WOS:001077246500001  
Indexed Date:  
2023-10-14  
Institution Address:  
China Agr Univ, Coll Engn, Beijing 100083, Peoples R China  
Institution Address:  
Tarim Univ, Coll Mech & Elect Engn, Alar 843300, Peoples R China  
Institution Address:  
Tarim Univ, Modern Agr Engn Key Lab Univ Educ Dept Xinjiang Uy, Alar 843300, Peoples R China  
Institution:  
Institution:  
Institution:

### 429. An ultra-high-definition multi-exposure image fusion method based on

摘要: Multiple exposure image fusion is a technique used to obtain high dynamic range images. Due to its low cost and high efficiency, it has received a lot of attention from researchers in recent years. Currently, most deep learning-based multiple exposure image fusion methods extract features from different exposure images using a single feature extraction method. Some methods simply rely on two different modules to directly extract features. However, this approach inevitably leads to the loss of some feature information during the feature extraction process, thus further affecting the performance of the model. To minimize the loss of feature information as much as possible, we propose an ultra-high-definition (UHD) multiple exposure image fusion method based on multi-scale feature extraction. The method adopts a U-shaped structure to construct the overall network model, which can fully exploit the feature information at different levels. Additionally, we construct a novel hybrid stacking paradigm to combine convolutional neural networks and Transformer modules. This combined module can extract both local texture features and global color features simultaneously. To more efficiently fuse and extract features, we also design a cross-layer feature fusion module, which can adaptively learn the correlation between features at different layers. Numerous quantitative and qualitative results demonstrate that our proposed method performs well in UHD multiple exposure image fusion.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1568-4946  
eISSN:  
1872-9681  
入藏号:  
WOS:001315316900001  
Indexed Date:  
2024-09-23  
Institution Address:  
Nanjing Univ Sci & Technol, Sch Comp Sci & Engn, Nanjing 210094, Peoples R China  
Institution Address:  
Guangxi Normal Univ, Guangxi Key Lab Multisource Informat Min & Secur, Guilin 541004, Peoples R China  
Institution Address:  
Nantong Univ, Sch Informat Sci & Technol, Nantong 226019, Peoples R China  
Institution:  
Institution:  
Institution:

### 430. Improved Deep Convolutional Neural Network-Based Method for Detecting

摘要: To address the issue of rapid and precise classification and recognition of winter jujube fruit in orchards, we propose an enhanced deep convolutional neural network and an improved loss function recognition method based on the YOLOv4 model. The multi -scale feature fusion Cross Stage Partial Connections (CSP) structure and Convolution Block Attention Module (CBAM) are used to enhance the model, minimise the network scale, amplify the feature extraction capability and improve the false detection in fruit classification as well as the missing detection of occlusions. The Softmax cross -entropy loss function replaced the Sigmoid binary cross -entropy loss function as the classification loss function. Additionally, the EIoU loss function was introduced to replace the CIoU loss function as the bounding box regression loss function. These changes were made to further reduce false detection in fruit classification and enhance the accuracy of prediction frames. The experimental results indicated that P (Precision), mAP (Mean Average Precision) and IoU (Intersection over Union) values for three classes of winter jujube fruits were 81.86%, 82.46% and 81.35%, respectively. The model has 26.9M parameters, a size of 108MB and a detection speed of 28.8 frames per second. This method boasts high accuracy and robustness, providing significant reference value for accurately detecting winter jujube fruit in orchards.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1816-093X  
eISSN:  
1816-0948  
入藏号:  
WOS:001180520900015  
Indexed Date:  
2024-04-12  
Institution Address:  
Baoding Univ, Coll Informat Engn, Baoding, Hebei, Peoples R China  
Institution Address:  
Hebei Agr Univ, Coll Informat Sci & Technol, Baoding, Hebei, Peoples R China  
Institution Address:  
Hebei Key Lab Agr Big Data, Baoding, Hebei, Peoples R China  
Institution Address:  
Shijiazhuang Univ, Coll Future Informat Technol, Shijiazhuang, Hebei, Peoples R China  
Institution:  
Institution:  
Institution:  
Hebei Key Lab Agr Big Data  
Institution:

### 431. Multi-scale fusion dehazing network for high-frequency information

摘要: At present, there is little work in the field of dehazing that introduces prior information into data-driven deep leaming methods, and most dehazing networks based on deep learning usually have high requirements on computer memory and computing power. To solve the above problems, this paper proposes a multi-scale fusion dehazing network for high-frequency information alignment (HFMS-Net). The network framework adoptes a eyeyle pattern: for the generator, residual connections are introduced at different depths of the lightweight convolutional neural network to make full use of the intermediate layer features of the network to achieve multi-scale feature fusion; for the discriminator, the network needs to extract texture information on its input to approximate the high-frequency information between the dehazed image and the hazy image, making the data-driven network more physically interpretable. Compared with PFDN, HEMS-Net achieves superior performance with about 1/5 of the memory footprint under the same setting, and the PSNR and SSIM are improved by 0.71 and 0.016, respectively. Through a large number of comparative experiments and ablation experiments, it is proved that the dehazing performance of this network has a certain improvement compared with the existing algorithms, and higher fidelity to texture information.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1007-2780  
eISSN:  
2097-3217  
入藏号:  
WOS:001043334500008  
Indexed Date:  
2023-09-08  
Institution Address:  
Shaanxi Univ Sci & Technol, Sch Elect Informat & Artificial Intelligence, Xian 710021, Peoples R China  
Institution:

### 432. Adjacent age classification algorithm of yellow-feathered chickens based

摘要: The age of yellow-feathered chicken is important to distinguish the freshness of meat quality in the trade of yellow-feathered chicken. To investigate whether the Convolutional Neural Network (CNN) model can be applied to the instar classification of yellow-feathered chickens, a multi-scale feature fusion model called Chicken\_Age\_Network (CANet) was proposed. The model uses Inception to construct a feature extraction layer and extract the feature information of each chicken face image to improve the classification accuracy of the model. First, the self-developed yellow-feathered chicken facial image collection application was used to collect images, and the yellow-feathered chicken image database was constructed by using day-age classification. Second, the standard face image of yellow-feathered chicken was obtained by using Structural Similarity Index Method detection(SSIM), image segmentation, background removal, and normalization. The adjacent age classification needs to extract more features, and CANet's feature extraction layer based on multi-scale feature fusion can extract features of different sizes. Another advantage of CANet is that the GMP (Global Max Pooling) replaces the final fully connected layer of the general CNN to reduce parameters and optimize the network model. Finally, chicken face images of adjacent days of age were tested and compared on VGG13, VGG19, DenseNet121, DenseNet161, SE-ResNet-20, MobileNet V1, ShuffleNet G2, ResNet50, ResNet34 and CSPDenseNet121. Test results show that CANet can quickly and accurately identify the age of yellow-feathered chickens. The classification accuracy on CANet model is 96.29 %, which is better than VGG13 (93.09 %), VGG19 (95.08 %), ResNet50 (85.53 %), ResNet34 (91.14 %), DenseNet121 (93.03 %), DenseNet161 (93.33 %), MobileNet V1(86.06 %), ShuffleNet G2(93.18 %), CSPDenseNet121(84.76 %) and SE-ResNet-20 (83.86 %). In order to verify the generalization of the model, experiments were carried out on the public data set CIFAR-10. Results show that CANet has good generalization and can be applied to other classification problems.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
被引频次合计:  
6  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:000888184900003  
Indexed Date:  
2022-12-14  
Institution Address:  
China Agr Univ, Coll Informat & Elect Engn, Beijing 100083, Peoples R China  
Institution Address:  
Guangdong Acad Agr Sci, Inst Anim Sci, Guangzhou 510640, Peoples R China  
Institution Address:  
State Key Lab Livestockand Poultry Breeding, Guangzhou 510640, Peoples R China  
Institution:  
Institution:  
Institution:  
State Key Lab Livestockand Poultry Breeding

### 433. Fabric defect detection method based on multi-scale fusion attention

摘要: Fabric defect detection is extremely important for the development of the textile industry, but the existing traditional image processing algorithms are not good enough to detect fabric defects, and the detection efficiency and accuracy of the classical deep learning model is not satisfactory, so this paper proposes an improved fabric defect detection method based on multi-scale fusion of attention mechanism YOLOv7-PCBS. Based on the YOLOv7 network structure, some of the standard convolutions of the backbone network are replaced with Partial Convolution modules, which reduces the amount of network computation and improves the network detection speed; add coordinate attention to enhance the ability of extracting the positional features of tiny defects in fabrics; reconfiguration of the SPPCSPC module to improve small target detection; optimization of Bidirectional Feature Pyramid Network (BiFPN) and design of Tiny- BiFPN for simple and fast multi-scale feature fusion; finally, a novel loss function SIoU with angular loss is introduced to facilitate the fitting of the true and predicted frames and enhance the accuracy of defect prediction. The results show that the algorithm achieves a mAP value of 94.4% on the detection of defects in solid-colored fabrics of six denim materials, which is an improvement of 15.1% compared to the original YOLOv7 algorithm, while the model achieves a frame rate of 59.5 per second. Compared with other traditional deep learning algorithms SSD and Faster-RCNN, the detection accuracies are improved by 21.6% and 15.2%, and the FPS values are improved by 78.1% and 101.0%, respectively. Therefore, the YOLOv7-PCBS fabric defect detection algorithm proposed in this paper makes the fabric defect detection results more accurate while realizing lightweight, which provides an important technical reference for the subsequent improvement of textile quality.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001358758800001  
Indexed Date:  
2024-11-26  
Institution Address:  
Nanjing Forestry Univ, Dept Mech, Nanjing 210037, Peoples R China  
Institution:

### 434. A Multi-Scale Rebar Detection Network with an Embedded Attention

摘要: With the rapid development of computer vision and machine vision, methods based on deep learning have achieved good results in the field of object detection, identification, and tracking. However, for the detection and identification of rebars in smart construction sites, it is very difficult to perform accurate real-time detection of rebars by using object detection technology on the equipment in the field because of the problems of the dense cross-section between bundled bars, the mutual adhesion of cross-section boundaries, and mutual occlusion between cross-sections. To address the above problems, we propose a multi-scale rebar detection network RebarNet with an embedded attention mechanism based on YOLOv5, combining the K-means++ algorithm, attention mechanism, a newly designed SD\_IoU Loss, and multi-scale feature fusion, aiming to solve the problems of missed and false detection in dense small object detection. Due to the problems of scarce rebar cross-section datasets, no publicly available large datasets, and weak rebar cross-sectional features, we constructed a new rebar cross-sectional dataset, used a semi-automatic annotation method to annotate part of the dataset, and then used the data enhancement algorithm to expand the rebar dataset. The experimental results show that the average accuracy (mAP) of our proposed RebarNet network is 97.9%, which is comparable to mainstream target detection algorithms such as Faster R-CNN, SSD, RetinaNet, CenterNet, CornerNet, YOLOv3, YOLOv4, and YOLOv5s. mAP0.5 is improved by 8.1%, 13%, 26.4%, 25.8%, 26.2%, 11.7%, 7.6%, and 9%, respectively. In addition, the frames per second (FPS) transmission reaches 89 frames per second, the model weight is only 17.0 MB. In summary, the proposed RebarNet can effectively reduce missed and false detections in the rebar counting detection task based on real-time detection.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
9  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
9  
eISSN:  
2076-3417  
入藏号:  
WOS:001037947400001  
Indexed Date:  
2023-08-10  
Institution Address:  
Henan Polytech Univ, Coll Comp Sci & Technol, Jiaozuo 454003, Peoples R China  
Institution:

### 435. A CNNA-Based Lightweight Multi-Scale Tomato Pest and Disease

摘要: Tomato is generally cultivated by transplanting seedlings in ridges and furrows. During growth, there are various types of tomato pests and diseases, making it challenging to identify them simultaneously. To address this issue, conventional convolutional neural networks have been investigated, but they have a large number of parameters and are time-consuming. In this paper, we proposed a lightweight multi-scale tomato pest and disease classification network, called CNNA. Firstly, we constructed a dataset of tomato diseases and pests consisting of 27,193 images with 18 categories. Then, we compressed and optimized the ConvNeXt-Tiny network structure to maintain accuracy while significantly reducing the number of parameters. In addition, we proposed a multi-scale feature fusion module to improve the feature extraction ability of the model for different spot sizes and pests, and we proposed a global channel attention mechanism to enhance the sensitivity of the network model to spot and pest features. Finally, the model was trained and deployed to the Jetson TX2 NX for inference of tomato pests and diseases in video stream data. The experimental results showed that the proposed CNNA model outperformed the pre-trained lightweight models such as MobileNetV3, MobileVit, and ShuffleNetV2 in terms of accuracy and all parameters, with a recognition accuracy of 98.96%. Meanwhile, the error rate, inference time for a single image, network parameters, FLOPs, and model size were only 1%, 47.35 ms, 0.37 M, 237.61 M, and 1.47 MB, respectively.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
Web of Science 核心合集中的 "被引频次":  
6  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
eISSN:  
2071-1050  
入藏号:  
WOS:001003452600001  
Indexed Date:  
2023-06-20  
Institution Address:  
JiLin Agr Univ, Coll Informat & Technol, Changchun 130118, Peoples R China  
Institution Address:  
Washington State Univ, Ctr Precis & Automated Agr Syst, Dept Biol Syst Engn, Prosser, WA 99350 USA  
Institution Address:  
Army Acad Armored Force, Non Commissioned Officer Sch, Changchun 130118, Peoples R China  
Institution:  
Institution:  
Institution:  
Army Acad Armored Force

### 436. A Face Detection and Standardized Mask-Wearing Recognition Algorithm

摘要: In the era of coronavirus disease (COVID-19), wearing a mask could effectively protect people from the risk of infection and largely reduce transmission in public places. To prevent the spread of the virus, instruments are needed in public places to monitor whether people are wearing masks, which has higher requirements for the accuracy and speed of detection algorithms. To meet the demand for high accuracy and real-time monitoring, we propose a single-stage approach based on YOLOv4 to identify the face and whether to regulate the wearing of masks. In this approach, we propose a new feature pyramidal network based on the attention mechanism to reduce the loss of object information that can be caused by sampling and pooling in convolutional neural networks. The network is able to deeply mine the feature map for spatial and communication factors, and the multi-scale feature fusion makes the feature map equipped with location and semantic information. Based on the complete intersection over union (CIoU), a penalty function based on the norm is proposed to improve positioning accuracy, which is more accurate at the detection of small objects; the new bounding box regression function is called Norm CIoU (NCIoU). This function is applicable to various object-detection bounding box regression tasks. A combination of the two functions to calculate the confidence loss is used to mitigate the problem of the algorithm bias towards determinating no objects in the image. Moreover, we provide a dataset for recognizing faces and masks (RFM) that includes 12,133 realistic images. The dataset contains three categories: face, standardized mask and non-standardized mask. Experiments conducted on the dataset demonstrate that the proposed approach achieves mAP@.5:.95 69.70% and AP75 73.80%, outperforming the compared methods.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
1424-8220  
入藏号:  
WOS:000997337000001  
PubMed ID:  
37430525  
Indexed Date:  
2023-06-09  
Institution Address:  
Chongqing Univ Post & Telecommun, Coll Automat, Chongqing 400065, Peoples R China  
Institution Address:  
Chongqing Univ Posts & Telecommun, Sch Sci, Chongqing 400065, Peoples R China  
Institution:  
Institution:

### 437. Apple recognition in complex environments based on FC-DETR

摘要: This study aims to address the challenges of apple recognition in complex environments by proposing a solution based on the FC-DETR model. Apple cultivation is an important part of agriculture in Xinjiang, but the increasing shortage of labor has driven the demand for automated harvesting technologies. Therefore, research and development of apple recognition technology in complex environments have become crucial. While CNN architecture models can accurately identify apples in conventional settings, their inherent receptive field limitations prevent them from fully capturing global features in noisy and complex environments, making it difficult to achieve the accuracy and robustness required for practical applications. To solve this problem, this study proposes a real-time object detection model, FC-DETR. This model incorporates the innovative FEMA-BasicBlock residual module, the CAFM cross-scale adaptive feature fusion module, and the novel Inner-WIoU loss function, thereby enhancing feature processing, multiscale feature selection and integration, and detection accuracy. Experimental results show that the FC-DETR model achieves an apple recognition accuracy of 87 % and a recall rate of 82 % in complex backgrounds while maintaining a lightweight design. This study not only makes significant advances in automated apple harvesting technology but also contributes to improving the efficiency and sustainability of the apple industry in Xinjiang and globally.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2405-8440  
入藏号:  
WOS:001315651700001  
PubMed ID:  
39974622  
Indexed Date:  
2024-09-25  
Institution Address:  
Tarim Univ, Coll Informat Engn, Minist Educ, Key Lab Tarim Oasis Agr, Alar, Peoples R China  
Institution:

### 438. New Progress in Intelligent Picking: Online Detection of Apple Maturity

摘要: In the realm of automated apple picking operations, the real-time monitoring of apple maturity and diameter characteristics is of paramount importance. Given the constraints associated with feature detection of apples in automated harvesting, this study proposes a machine vision-based methodology for the accurate identification of Fuji apples' maturity and diameter. Firstly, maturity level detection employed an improved YOLOv5s object detection model. The feature fusion section of the YOLOv5s network was optimized by introducing the cross-level partial network module VoVGSCSP and lightweight convolution GSConv. This optimization aimed to improve the model's multiscale feature information fusion ability while accelerating inference speed and reducing parameter count. Within the enhanced feature fusion network, a dual attention mechanism combining channel and spatial attention (GAM) was introduced to refine the color and texture feature information of apples and to increase spatial position feature weights. In terms of diameter determination, the contours of apples are obtained by integrating the dual features of color and depth images within the target boxes acquired using the maturity detection model. Subsequently, the actual area of the apple contour is determined by calculating the conversion relationship between pixel area and real area at the current depth value, thereby obtaining the diameter of the apples. Experimental results showed that the improved YOLOv5s model achieved an average maturity level detection precision of 98.7%. Particularly noteworthy was the detection accuracy for low maturity apples, reaching 97.4%, surpassing Faster R-CNN, Mask R-CNN, YOLOv7, and YOLOv5s models by 6.6%, 5.5%, 10.1%, and 11.0% with a real-time detection frame rate of 155 FPS. Diameter detection achieved a success rate of 93.3% with a real-time detection frame rate of 56 FPS and an average diameter deviation of 0.878 mm for 10 apple targets across three trials. Finally, the proposed method achieved an average precision of 98.7% for online detection of apple maturity level and 93.3% for fruit diameter features. The overall real-time inference speed was approximately 56 frames per second. These findings indicated that the method met the requirements of real-time mechanical harvesting operations, offering practical importance for the advancement of the apple industry.  
Web of Science 核心合集中的 "被引频次":  
10  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
eISSN:  
2073-4395  
入藏号:  
WOS:001210034800001  
Indexed Date:  
2024-05-04  
Institution Address:  
Shandong Agr Univ, Coll Mech & Elect Engn, Tai An 271002, Peoples R China  
Institution Address:  
Shandong Prov Key Lab Hort Machinery & Equipment, Tai An 271018, Peoples R China  
Institution Address:  
Shandong Acad Agr Machinery Sci, Jinan 250010, Peoples R China  
Institution Address:  
Shandong Agr Equipment Intelligent Engn Lab, Tai An 271002, Peoples R China  
Institution:  
Institution:  
Shandong Prov Key Lab Hort Machinery & Equipment  
Institution:  
Institution:  
Shandong Agr Equipment Intelligent Engn Lab

### 439. Research on Intelligent Oil Drilling Pipe Column Detection Method Based

摘要: The inadequate automation level in the transit of oil drilling tubular columns has led to significant inefficiencies and safety issues. To address these challenges, a real-time detection algorithm, ECS-YOLOv5s, has been proposed. This algorithm aims to improve the accuracy of drill pipe identification during operational processes, facilitating the automation of tubular column handling It has the potential to reduce drilling cycles and overall drilling costs. ECS-YOLOv5s enhance the detection accuracy of drill pipes by incorporating a Bidirectional Feature Pyramid Network (BiFPN) architecture with an improved multi-scale feature fusion network. The use of EfficientNet as the backbone network reduces the number of parameters and computations while effectively merging features from different layers. Additionally, the Spatial Pyramid Pooling (SPP) structure in the Neck is replaced with SPPF, and a Convolutional Block Attention Module (CBAM) is introduced to improve model robustness, reduce parameters and computations, and enhance the model's ability to detect dense targets. The ECS-YOLOv5s algorithm exhibits superior performance in drill pipe inspection, achieving a mean Average Precision (mAP) of 90.2%, a frame rate of 125 FPS, and a parameter count of only 37%. It achieves an accuracy of 98.6%, outperforming the original model by 9.2%. The comparative analysis demonstrates that the improved algorithm surpasses traditional models such as YOLOv5s, SSD, Faster-RCNN, and YOLOv7-tiny in both performance and accuracy. These findings provide valuable insights for the research on automated processing of tubular columns in intelligent oil drilling platforms.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001164137900001  
Indexed Date:  
2024-03-18  
Institution Address:  
Lanzhou Univ Technol, Sch Mech & Elect Engn, Lanzhou 730050, Peoples R China  
Institution:

### 440. DMF-Net: A Dual-Encoding Multi-Scale Fusion Network for Pavement Crack

摘要: Currently, cracks are the most common defect in pavement diseases. Long-term non-maintenance can lead to crack lengthening and expansion, causing serious traffic accidents, as well as shortening the service life of pavement cracks. Therefore, it is of utmost importance to maintain cracks at an early stage. Due to the effect of some challenging factors, such as various shape information of the cracks, complex textured backgrounds, light shadows, similar texture objects, micro cracks and other factors, accurate crack detection still faces a certain challenges. To solve the above problems, a dual-encoding multi-scale fusion network based on the combination of convolutional neural network (CNN) and transformer network is proposed, named DMF-Net. To obtain stronger feature representations, a dual-encoding path is built to acquire global context features and local detail information simultaneously, where global context features are extracted based on the transformer branch, and the local detail features are extracted based on the CNN branch to detect tiny details of the cracks. Meanwhile, an interactive attention learning (IAL) module is introduced to effectively fuse the global features from the transformer branch and the local detail information from the CNN branch, achieving mutual communication and learning of different feature information. In addition, to enrich the feature representation ability, an attention-based feature enhancement (AFE) module is introduced to acquire more global contexts. Furthermore, faced with the crack detection task with class imbalance issue, a triple attention module (TAM) is built to emphasize the micro cracks. Finally, in the segmentation prediction stage, the deep supervision mechanism is also introduced to accelerate the convergence speed of the model, and serve effective multi-scale feature fusion. Compared with the current mainstream segmentation models, excellent performance has been obtained, which could provide a feasible scheme for the early maintenance of pavement cracks. The source code about proposed DMF-Net is available at https://github.com/Bsl1/DMFNet.git.  
Web of Science 核心合集中的 "被引频次":  
12  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
13  
ISSN:  
1524-9050  
eISSN:  
1558-0016  
入藏号:  
WOS:001121634400001  
Indexed Date:  
2023-12-25  
Institution Address:  
Zhengzhou Univ, Sch Elect & Informat Engn, Zhengzhou 450001, Henan, Peoples R China  
Institution Address:  
Robot Percept & Control Engn Lab Henan Prov, Zhengzhou 450001, Henan, Peoples R China  
Institution Address:  
Edinburgh Napier Univ, Built Environm, Edinburgh EH10 5DT, Scotland  
Institution:  
Institution:  
Robot Percept & Control Engn Lab Henan Prov  
Institution:

### 441. A CNN-Transformer Hybrid Recognition Approach for sEMG-Based Dynamic

摘要: As a unique physiological electrical signal in the human body, surface electromyography (sEMG) signals always include human movement intention and muscle state. Through the collection of sEMG signals, different gestures can be effectively recognized. At present, the convolutional neural network (CNN) has been widely applied to different gesture recognition systems. However, due to its inherent limitations in global context feature extraction, it exists a certain shortcoming on high-precision prediction tasks. To solve this issue, a CNN-transformer hybrid recognition approach is proposed for high-precision dynamic gesture prediction. In addition, the continuous wavelet transform (CWT) is proposed for to acquire the time-frequency maps. To realize effective feature representation of local features from the time-frequency maps, an attention fusion block (AFB) is proposed to build the deep CNN network branch to effectively extract key channel information and spatial information from local features. Faced with the inherent limitations in global context feature extraction of CNNs, a transformer network branch is proposed to model the global relationship between pixels, called convolution and transformer (CAT) network branch. In addition, a multiscale feature attention (MFA) block is proposed for effective feature aggregation of local features and global contexts by learning adaptive multiscale features and suppressing irrelevant scale information. The experimental results on the established multichannel sEMG signal time-frequency map dataset show that the proposed CNN transformer hybrid recognition network has competitive recognition performance compared with other state-of-the-art recognition networks, and the average recognition speed of each spectrogram on the test set is only 14.7 ms. The proposed network can effectively improve network performance and identification efficiency.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
21  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
22  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001000275700017  
Indexed Date:  
2023-07-01  
Institution Address:  
Zhengzhou Univ, Sch Elect & Informat Engn, Zhengzhou 450001, Henan, Peoples R China  
Institution Address:  
Zhengzhou Univ, Sch Elect Engn, Zhengzhou 450001, Henan, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automat, Beijing 100190, Peoples R China  
Institution Address:  
Edinburgh Napier Univ, Sch Engn & Built Environm, Edinburgh EH10 5DT, Scotland  
Institution:  
Institution:  
Institution:  
Institution:

### 442. FCP-Net: A Feature-Compression-Pyramid Network Guided by Game-Theoretic

摘要: Medical image segmentation is a crucial step in diagnosis and analysis of diseases for clinical applications. Deep convolutional neural network methods such as DeepLabv3+ have successfully been applied for medical image segmentation, but multi-level features are seldom integrated seamlessly into different attention mechanisms, and few studies have fully explored the interactions between medical image segmentation and classification tasks. Herein, we propose a feature-compression-pyramid network (FCP-Net) guided by game-theoretic interactions with a hybrid loss function (HLF) for the medical image segmentation. The proposed approach consists of segmentation branch, classification branch and interaction branch. In the encoding stage, a new strategy is developed for the segmentation branch by applying three modules, e.g., embedded feature ensemble, dilated spatial mapping and channel attention (DSMCA), and branch layer fusion. These modules allow effective extraction of spatial information, efficient identification of spatial correlation among various features, and fully integration of multi-receptive field features from different branches. In the decoding stage, a DSMCA module and a multi-scale feature fusion module are used to establish multiple skip connections for enhancing fusion features. Classification and interaction branches are introduced to explore the potential benefits of the classification information task to the segmentation task. We further explore the interactions of segmentation and classification branches from a game theoretic view, and design an HLF. Based on this HLF, the segmentation, classification and interaction branches can collaboratively learn and teach each other throughout the training process, thus applying the conjoint information between the segmentation and classification tasks and improving the generalization performance. The proposed model has been evaluated using several datasets, including ISIC2017, ISIC2018, REFUGE, Kvasir-SEG, BUSI, and PH2, and the results prove its competitiveness compared with other state-of-the-art techniques.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
22  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
8  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
22  
ISSN:  
0278-0062  
eISSN:  
1558-254X  
入藏号:  
WOS:000804690300016  
PubMed ID:  
34982679  
Indexed Date:  
2022-06-13  
Institution Address:  
Hunan Univ, Coll Mech & Vehicle Engn, Engn Res Ctr Automot Elect & Control Technol, Changsha 410082, Peoples R China  
Institution Address:  
Northumbria Univ, Fac Engn & Environm, Newcastle Upon Tyne NE1 8ST, Tyne & Wear, England  
Institution:  
Institution:

### 443. YOLO-CIR: The network based on YOLO and ConvNeXt for infrared object

摘要: Transfer learning is widely used in infrared object detection algorithms, but these algorithms developed from visible usually ignored the characteristic of infrared images. In this paper, we propose a new object detection algorithm YOLO-CIR based on YOLO and ConvNext for infrared (IR) images. Specifically, to accommodate highbit-width infrared images, we propose the augmentation algorithm for infrared images and improve the preprocessing algorithm without bit-width compression. In addition, a multi-scale feature extraction network based on ConvNext was built to adapt the infrared images of low-resolution. Moreover, the coordinate attention module is introduced in the ConvNeXt block to focus on targets and suppress the background, and a split attention module is introduced in the neck to enhance feature fusion ability. The algorithm offers significant performance improvements over prevalent object detection algorithms, comparative experiment on the FLIR dataset shows that the algorithm outperforms the YOLOv5 by 3% and Faster R-CNN by 5.6% in map50 and has significant advantages in parameters and FLOPs.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
6  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
32  
被引频次合计:  
38  
ISSN:  
1350-4495  
eISSN:  
1879-0275  
入藏号:  
WOS:001007079300001  
Indexed Date:  
2023-06-26  
Institution Address:  
NORINCO Kunming Inst Phys, Nanjing Res Ctr, Nanjing 211106, Peoples R China  
Institution Address:  
Hohai Univ, Coll Sci, Nanjing 210024, Peoples R China  
Institution:  
NORINCO Kunming Inst Phys  
Institution:

### 444. DS-TransFusion:Automatic retinal vessel segmentation based on an

摘要: Retinal vascular segmentation holds significant value in medical research,playing an indispensable role in facilitating the screening of various diseases,such as diabetes,hypertension,and glaucoma.However,most current retinal vessel segmentation methods mainly rely on convolutional neural networks,which present limitations when dealing with long-term dependencies and global context connections.These limitations often result in poor segmentation of small blood vessels and low contrast between the ends of fundus blood vessel branches and the background.Addressing these issues is a pressing concern.To tackle these challenges,this paper proposes a new retinal blood vessel segmentation model,namely Dual Swin Transformer Fusion (DS-TransFusion).This model uses a two-scale encoder subnetwork based on a Swin Transformer,which is able to find correspondence and align features from heterogeneous inputs.Given an input image of a retinal blood vessel,the model first splits it into two nonoverlapping blocks of different sizes.These are then fed into the two branches of the encoder to extract coarse-grained and fine-grained features of the retinal blood vessels.At the jump j unction,DS-TransFusion introduces the Transformer interactive fusion attention (TIFA) module.The core of this module is to use a multiscale attention (MA) mechanism to facilitate efficient interaction between multiscale features.It integrates features from two branches at different scales,achieves effective feature fusion,enriches cross-view context modeling and semantic dependency,and captures long-term correlations between data from different image views.This,in turn,enhances segmentation performance.In addition,to integrate multiscale representation in the hierarchical backbone,DS-TransFusion introduces an MA module between the encoder and decoder.This module learns the feature dependencies across different scales,collects the global correspondence of multiscale feature representations,and further optimizes the segmentation effect of the model.The results showed that DS-TransFusion performed impressively on public data sets STARE,CHASEDB1,and DRIVE,with accuracies of 96.50%,97.22%,and 97.80%,and sensitivities of 84.10%,84.55%,and 83.17%,respectively.Experimental results show that DS-TransFusion can effectively improve the accuracy of retinal blood vessel segmentation and accurately segment small blood vessels.Overall,DS-TransFusion,as a Swin Transformer-based retinal vessel segmentation model,has achieved remarkable results in solving the problems of unclear segmentation of small vessels and global context connection.Experimental results on several public data sets have validated the effectiveness and superiority of this method,suggesting its potential to provide more accurate retinal vascular segmentation results for auxiliary screening of various diseases.  
摘要:  
视网膜血管的准确分割在辅助筛查各种疾病方面具有重大意义.然而,当前流行的模型仍存在细小血管的分割不清晰,以及眼底血管分支末端与背景的对比度较低等问题.针对这些问题,本文提出了一种全新的视网膜血管分割模型,命名为Dual Swin Transformer Fusion(DS-TransFusion).首先,DS-TransFusion采用基于Swin Transformer的双尺度编码器子网络,以提取视网膜血管的粗粒度和细粒度特征.其次,在跳跃连接处引入了Transformer交互融合注意力(TIFA)模块,用于丰富跨视图上下文建模和语义依赖,同时捕获来自不同图像视图的数据之间的长期相关性.最后,在编码器和解码器之间,DS-Trans- Fusion采用了多尺度注意力(MA),用于收集多尺度特征表示的全局对应关系,进一步优化模型的分割效果.实验结果表明,DSTransFusion在公共数据集STARE、CHASEDB1和DRIVE上表现出色,准确率分别达到了96.50%、97.22%和97.80%,灵敏度达到84.10%、84.55%和83.17%.实验表明DS-TransFusion能有效提高视网膜血管分割的精度,准确分割出细小血管.对视网膜血管分割的准确度、灵敏度和特异性都有大幅提高,与现有的SOTA方法相比具有更好的分割性能.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2095-9389  
入藏号:  
CSCD:7783688  
Indexed Date:  
2024-10-18  
Institution Address:  
辽宁工程技术大学软件学院, 葫芦岛, 辽宁 125105, 中国  
Institution Address:  
Software College,Liaoning Technical University, Huludao, Liaoning 125105, China  
Institution:  
辽宁工程技术大学软件学院  
Institution:  
Software College,Liaoning Technical University

### 445. High-performance deep transfer learning model with batch normalization

摘要: Deep learning and machine learning are cutting-edge methods for analysing images that have considerable potential. Artificial Neural Networks (A-NNs), one of the most well-known methods of computer intelligence, are now used in machine learning (ML) and deep transfer learning (DL) to raise plant production and quality. Identification and primary prevention of plant diseases at the appropriate time are essential for boosting productivity. Due to the phenomenon of minimally intense data in the background and foreground areas of the image, the extensive colour similarity between regions of unhealthy and normal leaves, the presence of noise in the sampling data, and changes in the location, size, and shape of plant leaf, it is difficult to correctly identify and classify plant diseases. In an effort to address these issues, a reliable technique for classifying plant diseases was developed by using a deep AlexNet CNN architecture as the main network with batch normalisation. In the three-step process, the first annotation is made to obtain the RoI (region of interest). The AlexNet CNN is therefore suggested for deep primary feature extraction in a constructed efficient network. The research demonstrates that the existing strategy is superior to more recent ones in terms of accuracy and dependability in recognising diseases in plants. Based on a deep transfer AlexNet CNN model, this research work developed a model for diseases identification and classification in plant leaves. It is trained using additional datasets that include a variety of plant leaf classifications and background images. From Plant Village and Kaggle, we gathered data on healthy and diseased tomato plant leaves. We are obtaining a near-balanced dataset containing ten different leaf disease kinds, such as bacterial, fungal, viral, and nutrient insufficiency. Ten classes have been considered for this research by gathering a dataset with associated images of the typical and abnormal tomato plant leaves. Considered in this work were the various labels for healthy and diseased tomato leaves, such as early blight, Bacterial spot, late bright mold, healthy, etc. Since deep CNN models have shown notable machine vision results, they are used in this case to diagnose and categorise plant illnesses from their leaves. As a result, the proposed CNN models can thus now be evaluated from confusion matrix using data analysis criteria, primarily focusing on metrics for evaluation like training and validation accuracy, loss, Recall, Precision, F1 score, processing speed, and performance.  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
2515-7620  
入藏号:  
WOS:001133064100001  
Indexed Date:  
2024-01-10  
Institution Address:  
K S Rangasamy Coll Technol, Dept Elect & Commun Engn, Namakkal, India  
Institution:

### 446. BED-YOLO: An Enhanced YOLOv8 for High-Precision Real-Time Bearing Defect

摘要: In industrial production, precise detection of bearing defects is crucial for optimal machinery performance and maintenance, directly impacting the efficiency of industrial systems and the field of instrumentation and measurement. To tackle the diverse types and unique characteristics of bearing defects, we introduce an approach for identifying defects in bearings, named bearing enhanced detection you only look once (BED-YOLO), which is based on convolutional neural networks (CNNs). We propose the intelligent feature concentration (IFC) module, a lightweight adaptive downsampling technique that exploits the attention mechanism to accurately control the feature compression process, prioritizing the retention of key features through the generation and normalization of spatial attention maps. Additionally, we design the efficient feature fusion for scalable convolution (EFFSC) module to capture and fuse multiscale features through convolution kernels of different sizes and optimize the computational efficiency using grouped convolution, which significantly improves the model expressiveness and processing speed. To ensure the robustness and reliability of our model, we conducted k-fold cross-validation on our BRG-dataset, which allowed us to thoroughly evaluate the model's performance and ensure its generalizability. The experimental results show that the BED-YOLO model demonstrates an excellent balance between performance and efficiency. The model achieves a mean average precision (mAP50) of 92.5%. Moreover, the model maintains high efficiency with a computational demand of only 7.7 GFLOPs and achieves processing speeds of 312.5 frames/s, while requiring only 2.5M parameters. These results highlight our model's superiority in speed and accuracy, making it particularly suitable for real-time applications that require rapid and precise detection, and well-equipped to meet the rigorous demands of industrial defect detection. Further tests on the MS COCO dataset underscore the model's remarkable adaptability and accuracy. Access to the methodology's code is provided through GitHub at https://github.com/YOLO-dennis/BED-YOLO.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001342557800016  
Indexed Date:  
2024-11-11  
Institution Address:  
Northeastern Univ, Coll Comp Sci & Engn, Shenyang 110819, Peoples R China  
Institution Address:  
Northeastern Univ, Coll Mech Engn & Automat, Shenyang 110819, Peoples R China  
Institution:  
Institution:

### 447. Remote Sensing Image Fusion With Task-Inspired Multiscale

摘要: Recently, convolutional neural networks (CNNs) have been developed for remote sensing image fusion (RSIF). To obtain competitive fusion performance, network design becomes more complicated by stacking convolutional layers deeper and wider. However, problems still remain when applying the existing networks in practical applications. On the one hand, researchers focus on improving spatial resolution but ignore that the fused images will be used in subsequent interpretation applications, e.g., objection detection. On the other hand, RSIF involves different tasks with different image sources, e.g., pansharpening of the panchromatic and multispectral image (MSI), hypersharpening of the panchromatic and hyperspectral image (HSI), and so on. However, the existing networks only solve one of them, failing to be compatible with other tasks. To address the above problems, a convenient task-inspired multiscale nonlocal-attention network (MNAN) is proposed for RSIF. The proposed MNAN focuses more on enhancing the multiscale targets in the scene when improving the resolution of the fused image. In addition, the proposed network can be applied to both pansharpening and hypersharpening tasks without any modification.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
10  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:000957555200001  
Indexed Date:  
2023-04-22  
Institution Address:  
Beijing Inst Technol, Sch Informat & Elect, Beijing 100081, Peoples R China  
Institution Address:  
Beijing Inst Technol, Beijing Key Lab Fract Signals & Syst, Beijing 100081, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Key Lab Network Informat Syst Technol NIST, Beijing, Peoples R China  
Institution Address:  
Univ Grenoble Alpes, LJK, INRIA, CNRS,Grenoble INP, Grenoble, France  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 448. AMFP-net: Adaptive multi-scale feature pyramid network for diagnosis of

摘要: Early detection of pneumoconiosis by routine health screening of workers in the mining industry is critical for preventing the progression of this incurable disease. Automated pneumoconiosis classification in chest X-ray images is challenging due to the low contrast of opacities, inter-class similarity, intra-class variation and the existence of artifacts. Compared to traditional methods, convolutional neural networks have shown significant improvement in pneumoconiosis classification tasks, however, accurate classification remains challenging due to mainly the inability to focus on semantically meaningful lesion opacities. Most existing networks focus on high level abstract information and ignore low level detailed object information. Different from natural images where an object occupies large space, the classification of pneumoconiosis depends on the density of small opacities inside the lung. To address this issue, we propose a novel two-stage adaptive multi-scale feature pyramid network called AMFP-Net for the diagnosis of pneumoconiosis from chest X-rays. The proposed model consists of 1) an adaptive multi-scale context block to extract rich contextual and discriminative information and 2) a weighted feature fusion module to effectively combine low level detailed and high level global semantic information. This two-stage network first segments the lungs to focus more on relevant regions by excluding irrelevant parts of the image, and then utilises the segmented lungs to classify pneumoconiosis into different categories. Extensive experiments on public and private datasets demonstrate that the proposed approach can outperform state-of-theart methods for both segmentation and classification.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0933-3657  
eISSN:  
1873-2860  
入藏号:  
WOS:001260709800001  
PubMed ID:  
38917599  
Indexed Date:  
2024-07-08  
Institution Address:  
Univ New South Wales, Sch Comp Sci & Engn, Sydney, NSW, Australia  
Institution Address:  
CSIRO DATA61, Sydney, Australia  
Institution:  
Institution:

### 449. MS-TCNet: An effective Transformer-CNN combined network using

摘要: Medical image segmentation is a fundamental research problem in the field of medical image processing. Recently, the Transformer have achieved highly competitive performance in computer vision. Therefore, many methods combining Transformer with convolutional neural networks (CNNs) have emerged for segmenting medical images. However, these methods cannot effectively capture the multi-scale features in medical images, even though texture and contextual information embedded in the multi-scale features are extremely beneficial for segmentation. To alleviate this limitation, we propose a novel Transformer-CNN combined network using multi-scale feature learning for three-dimensional (3D) medical image segmentation, which is called MS-TCNet. The proposed model utilizes a shunted Transformer and CNN to construct an encoder and pyramid decoder, allowing six different scale levels of feature learning. It captures multi-scale features with refinement at each scale level. Additionally, we propose a novel lightweight multi-scale feature fusion (MSFF) module that can fully fuse the different-scale semantic features generated by the pyramid decoder for each segmentation class, resulting in a more accurate segmentation output. We conducted experiments on three widely used 3D medical image segmentation datasets. The experimental results indicated that our method outperformed state-of-theart medical image segmentation methods, suggesting its effectiveness, robustness, and superiority. Meanwhile, our model has a smaller number of parameters and lower computational complexity than conventional 3D segmentation networks. The results confirmed that the model is capable of effective multi-scale feature learning and that the learned multi-scale features are useful for improving segmentation performance. We open-sourced our code, which can be found at https://github.com/AustinYuAo/MS-TCNet.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:001171279800001  
PubMed ID:  
38301516  
Indexed Date:  
2024-03-06  
Institution Address:  
Changchun Univ Sci & Technol, Sch Comp Sci & Technol, Changchun 130022, Peoples R China  
Institution Address:  
Changchun Univ Sci & Technol, Zhongshan Inst, Zhongshan 528437, Peoples R China  
Institution Address:  
Jilin Univ, Dept Hepatobiliary & Pancreat Surg, Hosp 1, Changchun 130061, Peoples R China  
Institution:  
Institution:  
Institution:

### 450. Remote sensing image object detection based on MobileViT and multiscale

摘要: A new algorithm is proposed based on MobileViT and multi-scale feature aggregation (referred to as FWMYOLOv7t) to address problems such as complex background interference, difficulty in extracting small objects, and object multi-scale differences in remote sensing image object detection. First, we design a multi-scale feature aggregation module to establish context dependencies for remote sensing targets, which improves the accuracy of detecting multiscale and small targets. Then, we utilize the MobileViT module to fuse the advantages of convolutional neural networks and vision transformers for effective local and global information encoding to suppress non-target noise interference. Finally, we introduce the Wise-IoU loss function, which focuses on ordinary quality anchor boxes to enhance the detection performance of the algorithm. Experimental evaluations on the public RSOD and NWPU VHR-10 dataset demonstrate that FWM-YOLOv7t can significantly improve the average accuracy of remote sensing image target detection. Furthermore, compared with other object detection algorithms, the FWM-YOLOv7t algorithm exhibits superior effectiveness in detecting complex, small, and multiscale objects in remote sensing imagery.  
摘要:  
针对遥感图像目标检测存在复杂背景干扰、微小目标提取难和目标多尺度差异问题,提出一种基于MobileViT和多尺度特征聚合的遥感图像目标检测算法(FWM-YOLOv7t)。首先设计多尺度特征聚合模块,建立遥感目标上下文依赖关系,提升多尺度目标和小目标检测精度;然后利用MobileViT模块,融合卷积神经网络和视觉Transformer优点,有效编码局部和全局信息,抑制非目标噪声干扰;最后引入Wise-IoU损失函数,重点关注普通质量锚框,提高算法检测性能。在公共数据集RSOD和NWPU VHR-10上的实验结果表明,FWMYOLOv7t能够显著提升遥感图像目标检测的平均准确率。与其他目标检测算法相比,FWM-YOLOv7t对复杂背景目标、小目标和多尺度目标的检测更有效。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1673-4785  
入藏号:  
CSCD:7852592  
Indexed Date:  
2025-02-09  
Institution Address:  
江西理工大学电气工程与自动化学院, 赣州, 江西 341000, 中国  
Institution Address:  
School of Electrical Engineering and Automation,Jiangxi University of Science and Technology, Ganzhou, Jiangxi 341000, China  
Institution:  
江西理工大学电气工程与自动化学院  
Institution:  
School of Electrical Engineering and Automation,Jiangxi University of Science and Technology

### 451. Multi-Contrast Complementary Learning for Accelerated MR Imaging

摘要: Thanks to its powerful ability to depict high-resolution anatomical information, magnetic resonance imaging (MRI) has become an essential non-invasive scanning technique in clinical practice. However, excessive acquisition time often leads to the degradation of image quality and psychological discomfort among subjects, hindering its further popularization. Besides reconstructing images from the undersampled protocol itself, multi-contrast MRI protocols bring promising solutions by leveraging additional morphological priors for the target modality. Nevertheless, previous multi-contrast techniques mainly adopt a simple fusion mechanism that inevitably ignores valuable knowledge. In this work, we propose a novel multi-contrast complementary information aggregation network named MCCA, aiming to exploit available complementary representations fully to reconstruct the undersampled modality. Specifically, a multi-scale feature fusion mechanism has been introduced to incorporate complementary-transferable knowledge into the target modality. Moreover, a hybrid convolution transformer block was developed to extract global-local context dependencies simultaneously, which combines the advantages of CNNs while maintaining the merits of Transformers. Compared to existing MRI reconstruction methods, the proposed method has demonstrated its superiority through extensive experiments on different datasets under different acceleration factors and undersampling patterns.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
5  
ISSN:  
2168-2194  
eISSN:  
2168-2208  
入藏号:  
WOS:001180907300007  
PubMed ID:  
38157466  
Indexed Date:  
2024-04-03  
Institution Address:  
Shandong Univ, Sch Informat Sci & Engn, Qingdao 266237, Peoples R China  
Institution Address:  
ASTAR, Inst High Performance Comp, Singapore 138632, Singapore  
Institution Address:  
Harbin Inst Technol Shenzhen, Shenzhen Key Lab Visual Object Detect & Recognit, Shenzhen 518055, Peoples R China  
Institution:  
Institution:  
Institution:

### 452. CTCNet: A CNN-Transformer Cooperation Network for Face Image

摘要: Recently, deep convolution neural networks (CNNs) steered face super-resolution methods have achieved great progress in restoring degraded facial details by joint training with facial priors. However, these methods have some obvious limitations. On the one hand, multi-task joint learning requires additional marking on the dataset, and the introduced prior network will significantly increase the computational cost of the model. On the other hand, the limited receptive field of CNN will reduce the fidelity and naturalness of the reconstructed facial images, resulting in suboptimal reconstructed images. In this work, we propose an efficient CNN-Transformer Cooperation Network (CTCNet) for face super-resolution tasks, which uses the multi-scale connected encoder-decoder architecture as the backbone. Specifically, we first devise a novel Local-Global Feature Cooperation Module (LGCM), which is composed of a Facial Structure Attention Unit (FSAU) and a Transformer block, to promote the consistency of local facial detail and global facial structure restoration simultaneously. Then, we design an efficient Feature Refinement Module (FRM) to enhance the encoded features. Finally, to further improve the restoration of fine facial details, we present a Multi-scale Feature Fusion Unit (MFFU) to adaptively fuse the features from different stages in the encoder procedure. Extensive evaluations on various datasets have assessed that the proposed CTCNet can outperform other state-of-the-art methods significantly. Source code will be available at https://github.com/IVIPLab/CTCNet.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
71  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
72  
ISSN:  
1057-7149  
eISSN:  
1941-0042  
入藏号:  
WOS:000964720100002  
PubMed ID:  
37030697  
Indexed Date:  
2023-05-09  
Institution Address:  
Nanjing Univ Posts & Telecommun, Inst Adv Technol, IVIP Lab, Nanjing 210046, Peoples R China  
Institution Address:  
Soochow Univ, Prov Key Lab Comp Informat Proc Technol, Suzhou 215006, Peoples R China  
Institution Address:  
Shanghai Univ, Sch Commun & Informat Engn, Shanghai 200444, Peoples R China  
Institution Address:  
Nanjing Univ Sci & Technol, Jiangsu Key Lab Image & Video Understanding Social, Nanjing 210094, Peoples R China  
Institution Address:  
Nanjing Univ Sci & Technol, Sch Comp Sci & Technol, Nanjing 210094, Peoples R China  
Institution Address:  
Chinese Univ Hong Kong, Ctr Math Artificial Intelligence, Dept Math, Hong Kong, Peoples R China  
Institution Address:  
Westlake Univ, Res Ctr Ind Future, Hangzhou 310024, Peoples R China  
Institution Address:  
Westlake Univ, Sch Engn, Hangzhou 310024, Peoples R China  
Institution Address:  
OPPO Res, Seattle, WA 98101 USA  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
OPPO Res

### 453. Detection and identification of tea leaf diseases based on AX-RetinaNet

摘要: The accurate detection and identification of tea leaf diseases are conducive to its precise prevention and control. Convolutional neural network (CNN) can automatically extract the features of diseased tea leaves in the images. However, tea leaf images taken in natural environments have problems, such as complex backgrounds, dense leaves, and large-scale changes. The existing CNNs have low accuracy in detecting and identifying tea leaf diseases. This study proposes an improved RetinaNet target detection and identification network, AX-RetinaNet, which is used for the automatic detection and identification of tea leaf diseases in natural scene images. AX-RetinaNet uses an improved multiscale feature fusion module of the X-module and adds a channel attention module, Attention. The feature fusion module of the X-module obtains feature maps with rich information through multiple fusions of multi-scale features. The attention module assigns a network adaptively optimized weight to each feature map channel so that the network can select more effective features and reduce the interference of redundant features. This study also uses data augmentation methods to solve the problem of insufficient samples. Experimental results show the detection and identification accuracy of AX-RetinaNet for tea leaf diseases in natural scene images is better than the existing target detection and identification networks, such as SSD, RetinaNet, YOLO-v3, YOLO-v4, Centernet, M2det, and EfficientNet. The AX-RetinaNet detection and identification results indicated the mAP value of 93.83% and the F1-score value of 0.954. Compared with the original network, the mAP value, recall value, and identification accuracy increased by nearly 4%, by 4%, and by nearly 1.5%, respectively.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
10  
Web of Science 核心合集中的 "被引频次":  
30  
被引频次合计:  
33  
ISSN:  
2045-2322  
入藏号:  
WOS:000757457000084  
PubMed ID:  
35140287  
Indexed Date:  
2022-03-04  
Institution Address:  
Anhui Univ, Natl Engn Res Ctr Agroecol Big Data Anal & Applic, Hefei, Anhui, Peoples R China  
Institution:

### 454. Multiscale Semantic Feature Optimization and Fusion Network for Building

摘要: Automatic building extraction has been applied in many domains. It is also a challenging problem because of the complex scenes and multiscale. Deep learning algorithms, especially fully convolutional neural networks (FCNs), have shown robust feature extraction ability than traditional remote sensing data processing methods. However, hierarchical features from encoders with a fixed receptive field perform weak ability to obtain global semantic information. Local features in multiscale subregions cannot construct contextual interdependence and correlation, especially for large-scale building areas, which probably causes fragmentary extraction results due to intra-class feature variability. In addition, low-level features have accurate and fine-grained spatial information for tiny building structures but lack refinement and selection, and the semantic gap of across-level features is not conducive to feature fusion. To address the above problems, this paper proposes an FCN framework based on the residual network and provides the training pattern for multi-modal data combining the advantage of high-resolution aerial images and LiDAR data for building extraction. Two novel modules have been proposed for the optimization and integration of multiscale and across-level features. In particular, a multiscale context optimization module is designed to adaptively generate the feature representations for different subregions and effectively aggregate global context. A semantic guided spatial attention mechanism is introduced to refine shallow features and alleviate the semantic gap. Finally, hierarchical features are fused via the feature pyramid network. Compared with other state-of-the-art methods, experimental results demonstrate superior performance with 93.19 IoU, 97.56 OA on WHU datasets and 94.72 IoU, 97.84 OA on the Boston dataset, which shows that the proposed network can improve accuracy and achieve better performance for building extraction.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
16  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
16  
eISSN:  
2072-4292  
入藏号:  
WOS:000671021900001  
Indexed Date:  
2021-07-19  
Institution Address:  
Univ Putra Malaysia UPM, Fac Engn, Dept Civil Engn, Serdang 43400, Selangor, Malaysia  
Institution Address:  
Univ Putra Malaysia UPM, Fac Engn, Geospatial Informat Sci Res Ctr GISRC, Serdang 43400, Selangor, Malaysia  
Institution Address:  
Panzhihua Univ, Sch Civil & Architecture Engn, Dept Geog Informat, Panzhihua 617000, Peoples R China  
Institution Address:  
Panzhihua Univ, Sch Civil & Architecture Engn, Remote Sensing Res Ctr, Panzhihua 617000, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 455. Segmentation-assisted classification model with convolutional neural

摘要: Detecting weld defects in battery trays is crucial for the safety of new energy vehicles. Existing methods for weld surface defect detection, relying on traditional computer vision algorithms and convolutional neural networks with substantial image-level labeled data, face challenges in accurately identifying small defects, especially with limited samples. To address these issues, we developed an innovative Segmentation-Assisted Classification with Convolutional Neural Networks (SACNN) model. SACNN integrates a common feature extraction subnet, a segmentation subnet enhanced by a multi-scale feature fusion module, and a classification subnet specifically designed for precise defect detection. A joint loss function co-trains the segmentation and classification subnets using both image-level and pixel-level labels, enhancing the model's ability to accurately detect small defect regions. Our model demonstrates notable improvement, achieving accuracy gains ranging from 2% to 18% compared to existing state-of-the-art methods, with an overall accuracy of 94.09% on an industrial dataset of battery tray welds. To further evaluate the generalization capability of our model, we evaluated it on the publicly available Magnetic Tile dataset, achieving state-of-the-art results in this challenging context. Additionally, we conducted comprehensive ablation studies to validate the contribution of each component in our approach and utilized visualization techniques to enhance the interpretability of our model. These advancements represent a significant contribution to the state of the art in aluminum alloy weld defect detection.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0965-9978  
eISSN:  
1873-5339  
入藏号:  
WOS:001333951700001  
Indexed Date:  
2024-10-23  
Institution Address:  
Sichuan Univ, Sch Mech Engn, Chengdu 610065, Peoples R China  
Institution Address:  
Sichuan Univ, Yibin Ind Technol Res Inst, Yibin 644005, Peoples R China  
Institution:  
Institution:

### 456. YOLO-SM: A Lightweight Single-Class Multi-Deformation Object Detection

摘要: Recently, object detection witnessed vast progress with the rapid development of Convolutional Neural Networks (CNNs). However, object detection is mainly for multi-class tasks, and few networks are used to detect single-class multi-deformation objects. This paper aims to develop a lightweight object detection network for single-class multi-deformation objects to promote the practical application of object detection networks. First, we design a Densely Connected Multi-scale (DCM) module to augment the semantic information extraction of deformation objects. With the DCM module and other strategies incorporated, we design a lightweight backbone structure for object detection, namely, DCMNet. Then, we construct a lightweight Neck structure Ghost Multi-scale Feature (GMF) module for feature fusion using a feature linear generation strategy. Finally, with the DCMNet and GMF module, we propose the object detection network YOLO-SM for single-class multi-deformation objects. Extensive experiments demonstrate that our proposed backbone structure, DCMNet, significantly outperforms the state-of-the-art models. YOLO-SM achieves 97.66% mean Average Precision ($mAP$) on the Barcode public dataset, which is higher than other state-of-the-art object detection models, and achieves an inference time of 55.45 frames per second (FPS), proving that the YOLO-SM has a good performance tradeoff between speed and accuracy in detecting single-class multi-deformation objects. Furthermore, in the single-class multi-deformation Crack public dataset, the $mAP$ of 86.11% is achieved, and an $mAP$ of 99.84% is obtained in the multi-class dataset Dish20, which is much higher than other state-of-the-art object detection models, proving that the YOLO-SM has good generalization ability.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
6  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
2471-285X  
入藏号:  
WOS:001181499400001  
Indexed Date:  
2024-03-24  
Institution Address:  
Ritsumeikan Univ, Res Org Sci & Technol, Kusatsu 5258577, Japan  
Institution Address:  
Ritsumeikan Univ, Coll Sci & Engn, Kusatsu 5258577, Japan  
Institution:  
Institution:

### 457. Pavement crack detection method based on multi-scale feature enhancement

摘要: To solve the problems of incomplete pavement crack detection and discontinuous segmentation,a detection network MFENet for pavement cracks based on multi-scale feature enhancement was proposed,and the detection,classification and segmentation of end-to-end pavement crack images were realized.A multi-scale attention-based feature enhancement module was designed,and the mapping relationships of the weight coefficients of the upper multi-scale feature channels with those of the lower feature channels in the network model were determined to highlight the feature outputs from the effective channels.Based on the correlation between the coordinate information of the pavement crack and the semantic information of the pixels in physical location,a multi-semantic feature correlation module was designed and thereby feature fusion and enhancement among different semantic information were achieved.Then,the foreground features of the pavement crack image were filtered by feature dimension transformation.A quantitative evaluation method for deep feature intensity was proposed to improve the interpretability of the model's feature extraction ability.Research results on selfcollected dataset show that the average precision and average recall of the MFENet in pavement crack image detection are 4.3%and 5.4% higher than those of the Mask R-CNN,respectively, and 14.6% and 14.3% higher than those of the baseline model RDSNet,respectively.The average precision and average recall of the MFENet in pavement crack image segmentation are 6.6%and 8.8% higher than those of the Mask R-CNN,respectively,and 8.1% and 9.7% higher than those of the RDSNet,respectively.In the comparison with the Mask R-CNN and other mainstream methods,the images of different types of pavement cracks are detected and segmented with the highest accuracy by the MFENet.Research results on public datasets(CFD and CRACK500)show that the detection and segmentation accuracy of the MFENet are invariably higher than those of the Mask R-CNN and other mainstream methods on the datasets covering different scenarios,indicating the higher robustness of the proposed method.In addition,the processing speed of the MFENet is also faster than that of the RDSNet on different datasets.  
摘要:  
针对路面裂缝检测不完整和分割出现断裂的问题,提出了一种多尺度特征增强的路面裂缝检测网络MFENet,实现端到端的路面裂缝图像检测、分类和分割处理;设计了多尺度注意力特征增强模块,建立了网络模型的上层多尺度特征通道与底层特征通道权重系数之间的映射关系,以提升有效通道的特征输出;基于路面裂缝的坐标信息和像素语义信息在物理位置上的相关性,设计了多语义特征关联模块,实现不同语义信息之间的特征融合增强,并通过特征维度转换实现对路面裂缝图像的前景特征过滤;提出了一种针对深度特征强度进行量化评估的方法,用于提升模型提取特征能力的可解释性。在自采集数据集上的研究结果表明:MFENet对路面裂缝图像检测的平均精准率和平均召回率相比Mask R-CNN分别提升了4.3%和5.4%,相比基线模型RDSNet分别提升了14.6%和14.3%;MFENet对路面裂缝图像分割的平均精准率和平均召回率相比Mask RCNN分别提升了6.6%和8.8%,相比RDSNet分别提升了8.1%和9.7%;与Mask R-CNN等主流方法相比,MFENet对不同类型路面裂缝图像的检测、分割精度最高。在公开数据集(CFD、 CRACK500)上的研究结果表明:在不同场景下的数据集上,MFENet的检测、分割精度均高于Mask R-CNN等主流方法,模型的鲁棒性更强。另外与RDSNet相比,MFENet在不同数据集上的处理速度也均有所提升。  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1671-1637  
入藏号:  
CSCD:7411409  
Indexed Date:  
2023-06-10  
Institution Address:  
长安大学信息工程学院, 西安, 陕西 710064, 中国  
Institution Address:  
东南大学交通学院, 南京, 江苏 211189, 中国  
Institution Address:  
School of Information Engineering,Chang'an University, Xi'an, Shaanxi 710064, China  
Institution Address:  
School of Transportation,Southeast University, Nanjing, Jiangsu 211189, China  
Institution:  
长安大学信息工程学院  
Institution:  
东南大学交通学院  
Institution:  
School of Information Engineering,Chang'an University  
Institution:  
School of Transportation,Southeast University

### 458. Multi-view Remote Sensing Image Scene Classification by Fusing

摘要: Objectives: Remote sensing scene classification provides new possibilities for the application of high-resolution images, and how to effectively realize scene recognition from high-resolution remote sensing images is still an important challenge. The existing scene classification methods only use remote sensing images from one viewpoint for scene recognition, which cannot accurately express the semantic information of complex high-resolution remote sensing images, and the accuracy of scene classification is difficult to be further improved. Methods: To solve this problem, a multi-view scene classification method for remote sensing images is proposed. First, the aerial image and ground image are constructed into a positive and negative image pair, and divided into training dataset, validation dataset and test dataset. Second, a convolutional neural network with fusion multi-scale attention is constructed, and features with fusion attention and strong representation ability are obtained through feature fusion module, so as to integrate different feature information and realize multi-scale feature learning. Third, the trained multi-scale attention network is used to extract features from aerial image and ground image,respectively. Finally, the fused features are used to classify scenes based on the fused features using support vector machine. To demonstrate the performance of the proposed multi-scale attention network, we conduct experiments on two publicly available benchmark datasets-the AiRound and the CV-BrCT datasets. Results: The proposed method achieves remarkable performance, with the highest accuracy of 93.13% in the AiRound dataset and 85.18% in the CV-BrCT dataset, which improves the accuracy of single-view scene classification. Conclusions: The results demonstrate that the complementary information provided by multi-view images can further improve the performance of remote sensing scene classification.  
摘要:  
针对现有场景分类方法特征表征能力差以及单视角遥感影像分类精度难以提升的问题,提出一种融合多尺度注意力的多视角遥感影像场景分类方法。首先,将航空图像和地面图像构造成正负图像对,并划分为训练集、验证集和测试集;其次,构建融合多尺度注意力的卷积神经网络并训练,通过特征融合模块得到融合注意力且表征能力更强的特征,实现多尺度特征学习;然后,利用训练的多尺度注意力网络分别提取航空图像和地面图像特征并进行融合;最后,基于融合后的特征使用支持向量机进行场景分类。实验结果表明,相比现有方法,所提方法在两个公开数据集上均取得了更高的分类精度,改善了单视角场景分类效果,同时也证明了多视角所提供的补充信息能进一步提升遥感场景分类的准确性。  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
1671-8860  
入藏号:  
CSCD:7665194  
Indexed Date:  
2024-05-17  
Institution Address:  
南京信息工程大学遥感与测绘工程学院, 南京, 江苏 210044, 中国  
Institution Address:  
南京信息工程大学遥感与测绘工程学院;;北京师范大学, ;;遥感科学国家重点实验室, 南京;;, 江苏;;北京 210044;;100875, 中国  
Institution Address:  
武汉大学, 测绘遥感信息工程国家重点实验室, 武汉, 湖北 430079, 中国  
Institution Address:  
School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China  
Institution Address:  
School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology;;Beijing Normal University, ;;State Key Laboratory of Remote Sensing Science, Nanjing;;, ;;Beijing 210044;;100875  
Institution Address:  
Mapping and Remote Sensing, Wuhan University, State Key Laboratory of Information Engineering in Surveying, Wuhan, Hubei 430079, China  
Institution:  
南京信息工程大学遥感与测绘工程学院  
Institution:  
南京信息工程大学遥感与测绘工程学院;;北京师范大学  
Institution:  
武汉大学  
Institution:  
School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology  
Institution:  
School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology;;Beijing Normal University  
Institution:  
Mapping and Remote Sensing, Wuhan University

### 459. Visualization for Explanation of Deep Learning-Based Fault Diagnosis

摘要: Permanent magnet synchronous motor (PMSM) is widely used in various production processes because of its high efficiency, fast reaction time, and high power density. With the continuous promotion of new energy vehicles, timely detection of PMSM faults can significantly reduce the accident rate of new energy vehicles, further enhance consumers' trust in their safety, and thus promote their popularity. Existing fault diagnosis methods based on deep learning can only distinguish different PMSM faults and cannot interpret and analyze them. Convolutional neural networks (CNN) show remarkable accuracy in image data analysis. However, due to the "black box" problem in deep learning models, the diagnostic results regarding providing accurate information to the user are uncertain. This paper proposes a motor fault diagnosis method based on improved deep residual network (ResNet) and gradient-weighted class activation mapping (Grad-CAM) to analyze demagnetization and eccentricity faults of permanent magnet synchronous motors, and the uncertainty limitation of fault diagnosis based on the convolutional neural network is overcome by the visual interpretation method. The improved ResNet is formed by using ResNet9 as the backbone network, replacing the last convolution layer with a atrous spatial pyramid pooling (ASPP), and adding a multi-scale feature fusion module and attention channel mechanism (CAM). The proposed model not only retains the effective extraction of image features by ResNet9 but also enhances the sensitivity field of the network through the hollow convolution pyramid and realizes the feature fusion of the web on different scales through the multi-scale feature fusion module (MSFFM), further improving the diagnostic accuracy of the network on different types of fault features. The diagnostic effect of the network is verified on the selfmade data set, which mainly includes five states: normal (He), 25% demagnetization (De25), 50% demagnetization (De50), 10% static eccentricity (Se10), and 20% static eccentricity (Se20). The number of pictures in the training set is 6000, and the number in the test set is 1500. The average diagnostic accuracy of the improved ResNet on this dataset is 99.00%, which is 1.04%, 8.89%, 4.58%, and 7.22% higher than that of the multi-column convolutional neural network (MCNN), Bi-directional long short-term memory (Bi-LSTM), deep belief network (DBN), and recurrent neural network (RNN) models, respectively. Finally, gradient activation heat maps were used to globally average pool the final output feature map of the network to obtain feature weights. They were superimposed with the original image to get gradient activation heat maps of different grayscale images. The warmer the tone of the heat map, the greater the impact on the network diagnosis results, and then the demagnetization and eccentricity fault characteristics of the permanent magnet synchronous motor were determined-visual characterization of quantitative analysis.  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001126979000026  
Indexed Date:  
2024-01-11  
Institution Address:  
Guizhou Univ, Sch Elect Engn, Guiyang 550025, Peoples R China  
Institution:

### 460. Apple disease identification using improved Faster R-CNN

摘要: To analyze apple diseased leaf images,it is difficult to locate and identify these diseased leaves with small scale lesion and complex background in the actual application scenarios.In this study,five apple leaf diseases,i.e.,alternaria leaf spot,apple scab,gray spot,cedar rust and mosaic were investigated,and an improved Faster R-CNN based apple diseased leaf detection method was proposed.Firstly,the training set data was expanded through the data augmentation operation (including rotating,random brightness enhancement,random chromaticity enhancement,random contrast enhancement and sharpening) to enhance the robustness of the model.Then the augmented training set images were trained through the improved Faster R-CNN to make the detection model more reliable.For the improved Faster R-CNN model,the attention separation mechanism ResNest (split-attention networks) was adopted as the backbone to make the model focusing on the more useful information to enhance the feature extraction ability according to the feature representation through the weighted combination.To enhance the robustness of the feature information and improve the generalization ability,a feature pyramid network (FPN) was added to fuse multi-scale features,which effectively used the deep and shallow features of the network.Meanwhile,the cascade mechanism was adopted to optimize the generation mechanism of the suggestion box,so that the detection box location was more accurate.The mAP (mean average precision) of the improved model reached 86.2%,which is 8.7% higher than that of the previous Faster R-CNN model.The accuracy of the model reached 98.3% and the average detection time of the model was 0.092 s,which can effectively identify apple leaf lesions.The experimental results showed that the improved Faster R-CNN model could accurately and quickly identify small target lesions of apple leaves and lesions under complex background,and improve the accuracy of model recognition.The images in the data set included picked leaves and non-picked leaves,so this method could realize the nondestructive identification of apple leaf diseases,which provided a scientific basis for the early detection and prevention of apple diseases.  
摘要:  
针对苹果叶片图像中小尺度病斑和复杂背景带来的病斑目标难以精确定位和识别的问题,以苹果的斑点落叶病、黑星病、灰斑病、雪松锈病和花叶病为研究对象,提出一种基于改进Faster R-CNN的苹果叶片病害识别方法。先通过数据增广操作对训练集数据进行扩充以增强模型鲁棒性,再通过对增广训练集图像进行训练来得到一个可靠的病害识别模型。改进后的模型使用拆分注意力网络(ResNest)作为骨干特征提取网络,使模型更加关注对提升病斑检测性能有用的信息,以增强模型对特征的提取能力;通过添加特征金字塔网络(FPN)进行多尺度特征融合,以增强特征信息的鲁棒性,提高模型的泛化能力;采用级联机制对建议框生成机制进行优化,使检测框定位更加准确。改进后的Faster R-CNN模型的平均精度均值(mAP)达到86.2%,与改进前相比,其平均精度提升了8.7%,对单张病害图像的识别准确率达到98.3%,单张图像平均检测时间0.092 s,能有效识别苹果叶片病斑。实验结果表明,改进后的Faster R-CNN模型能准确快速地实现对苹果叶片小目标病斑和复杂背景下病斑的识别,提升模型识别的精准度。该识别方法可在实际场景下使用,无须特意采摘叶片实现对苹果叶片病害的无损测量识别,可为苹果病害的早期干预和治疗提供科学依据。  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
7  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
10  
ISSN:  
2096-1359  
入藏号:  
CSCD:7152943  
Indexed Date:  
2022-04-14  
Institution Address:  
山东农业大学信息科学与工程学院, 泰安, 山东 271018, 中国  
Institution Address:  
泰山学院数学与统计学院, 泰安, 山东 271000, 中国  
Institution Address:  
College of Information Science and Engineering,Shandong Agricultural University, Tai'an, Shandong 271018, China  
Institution Address:  
School of Mathematics and Statistics,Taishan University, Tai'an, Shandong 271000, China  
Institution:  
山东农业大学信息科学与工程学院  
Institution:  
泰山学院数学与统计学院  
Institution:  
College of Information Science and Engineering,Shandong Agricultural University  
Institution:  
School of Mathematics and Statistics,Taishan University

### 461. A Novel Arcing Detection Model of PantographCatenary for High-Speed

摘要: The pantograph-catenary system is crucial for transferring electrical power from catenary lines to electrified train and the occurrence of arcing could damage railway operations; thus, it is important to detect arcing. Detecting arcing in complex scenes and detecting different sizes and shapes of arcing is still a challenge. To overcome these issues, a robust image-based semantic segmentation model named arc multiscale fusion (ArcMSF) is proposed for arcing detection of pantograph-catenary, which designs a novel hybrid multiscale feature fusion model that aggregates Transformer with convolutional neural network (CNN) to realize arcing pixel segmentation. A down-top decoder for combining low-level features with high-level features is designed to achieve multiscale-level arcing feature detection in complex scenes. Inspired by the arcing image properties that arcing is always the brightest, the global max features and global threshold features are designed to augment the arcing features. Experiments on IVAIS-PCA2021 dataset and comparative experiments are conducted to demonstrate the effectiveness of the ArcMSF, which can achieve an 89.13% segmentation accuracy and a fast inference speed of 23.84 ms. Moreover, the detection results have a clear depiction of edge details.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000981235500013  
Indexed Date:  
2023-06-01  
Institution Address:  
Southwest Jiaotong Univ, Sch Elect Engn, Chengdu 610097, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, Sch Mech Engn, Chengdu 610097, Peoples R China  
Institution:  
Institution:

### 462. An Ultrasmall Bolt Defect Detection Method for Transmission Line

摘要: Bolt defect inspection is an important work in transmission line inspection. Due to the small size of bolts in the transmission line inspection images, existing algorithms are difficult to extract valuable features and achieve poor performance on bolt defect detection. This article proposed an ultrasmall bolt defect detection model (UBDDM) based on a deep convolutional neural network (DCNN), including an ultrasmall object perception module (UOPM) and a local bolt detection module (LBDM). In this article, UOPM is first constructed to realize coarse region recognition for the salient region of bolts in the inspection images, and the high-resolution image blocks are obtained from the original image according to the recognition results. Then, LBDM is constructed to intelligently identify the bolt defects from the high-resolution image blocks. Considering that the features of ultrasmall targets are difficult to extract, feature extraction networks are constructed based on residual network (ResNet)-50, and the hybrid attention mechanism and multiscale feature fusion are introduced to further improve the network's ability to extract shallow features. This method uses two-stage detection to realize end-to-end bolt defect detection but only needs to provide a single-stage target detection label, which greatly reduces the workload of data labeling. Experimental results show that the proposed method achieves excellent performance on bolt defect detection in inspection images.  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
43  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
45  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000936149000002  
Indexed Date:  
2023-03-16  
Institution Address:  
Wuhan Univ, Sch Elect Engn & Automat, Wuhan 430072, Peoples R China  
Institution Address:  
Univ Denver, Elect & Engn Dept, Denver, CO 80208 USA  
Institution:  
Institution:

### 463. ACA-Net: An Adaptive Convolution and Anchor Network for Metallic Surface

摘要: Metallic surface defect detection is critical to ensure the quality of industrial products. Recently, human-advanced surface defect detection algorithms have been proposed. Most of these algorithms rely on convolutional neural networks (CNN) and an anchoring scheme. However, a convolution unit only samples the input feature maps at fixed shapes and locations. Similarly, a set of anchors are uniformly predefined with fixed scales and shapes, which increases the difficulties of bounding box regression. Therefore, we propose an adaptive convolution and anchor network for metallic surface defect detection, named ACA-Net. Specifically, an adaptive convolution and anchor (ACA) module is proposed, which mainly consists of adaptive convolution and an adaptive anchor. Firstly, an adaptive convolution module (ACM) is designed, which adaptively determines the location and shape of each convolution unit. In addition, a multi-scale feature adaptive fusion (MFAF) is proposed, which is used in ACM to extract and integrate multi-scale features. Then, an adaptive anchor module (AAM) is proposed to yield more suitable anchor boxes by adaptively adjusting shapes. Extensive experiments on NEU-DET dataset and GC10 dataset validate the performance of the proposed approach. ACA-Net achieves 1.8% on NEU-DET dataset higher Average Precision (AP) than GA-RetinaNet. Furthermore, the proposed ACA module is also adopted in GA-Faster R-CNN, improving the AP by 1.2% on NEU-DET dataset.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
eISSN:  
2076-3417  
入藏号:  
WOS:000846248400001  
Indexed Date:  
2022-09-04  
Institution Address:  
Henan Univ Technol, Sch Mech & Elect Engn, Zhengzhou 450001, Peoples R China  
Institution Address:  
Henan Int Joint Lab Grain Informat Proc, Zhengzhou 450001, Peoples R China  
Institution Address:  
Henan Univ Technol, Coll Informat Sci & Engn, Zhengzhou 450001, Peoples R China  
Institution:  
Institution:  
Henan Int Joint Lab Grain Informat Proc  
Institution:

### 464. Toward Real Hyperspectral Image Stripe Removal via Direction Constraint

摘要: In hyperspectral imaging (HSI), stripe noise is one of the most common noise types that adversely affects its application. Convolutional neural networks (CNNs) have contributed to state-of-the-art performance in HSI destriping given their powerful feature extraction and learning capabilities. However, it is difficult to obtain paired training samples for real data. Most CNN destriping methods construct a paired training dataset with simulated stripe noise for network training. However, when the stripe noise of real data is complex, destriping performance of the model is constrained. To solve this problem, this study proposes a real HSI stripe removal method using a toward real HSI stripe removal via direction constraint hierarchical feature cascade network (TRS-DCHC). TRS-DCHC uses the stripe noise extract subnetwork to extract stripe patterns from real stripe-containing HSI data and incorporates clean images to form paired training samples. The destriping subnetwork advantageously utilizes a wavelet transform to explicitly decompose stripe and stripe-free components. It also adopts multi-scale feature dense connections and feature fusion to enrich feature information and deeply mine the discriminate features of stripe and stripe-free components. Our experiments on both simulated and real data of various loads showed that TRS-DCHC features better performance in both simulated and real data compared with state-of-the-art method.  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
eISSN:  
2072-4292  
入藏号:  
WOS:000820988700001  
Indexed Date:  
2022-07-13  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & R, Wuhan 430079, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Remote Sensing & Informat Engn, Wuhan 430079, Peoples R China  
Institution:  
Institution:

### 465. Bearing life prediction based on multi-scale features and attention

摘要: A bearing RUL prediction method based on multi-scale features and attention mechanism was proposed aiming at the problem that the previous remaining useful life (RUL) prediction methods were insufficient in mining bearing degradation information and ignored the difference in the contribution of different features,which affected the prediction accuracy.Several time-domain and frequency-domain features of the original bearing vibration signal at multiple scales were calculated as the input feature set.The multi-scale feature set was input into the network,and the attention module was used to adaptively assign the best weights to different features.Then the convolutional neural network (CNN) module was used for deep feature extraction and multi-scale feature fusion.The RUL prediction value was obtained through the feedforward neural network (FNN) module mapping.The proposed method was applied to the public bearing datasets for comparative studies.Results showed the superior prediction performance of the proposed method.  
摘要:  
针对以往剩余使用寿命(RUL)预测方法对轴承退化信息挖掘不充分、忽视不同特征贡献度差异,影响预测准确性的问题,提出基于多尺度特征与注意力机制的轴承RUL预测方法.在多个尺度下计算轴承原始振动信号的若干时域和频域特征,作为输入特征集.将多尺度特征集输入到网络中,以注意力模块为不同特征自适应地分配最佳权重,以卷积神经网络(CNN)模块进行深层特征提取与多尺度特征融合,通过前馈神经网络(FNN)模块映射得到RUL预测值.通过公开的轴承数据集进行实验验证,与其他RUL预测方法相比,所提方法的预测性能更优越.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1008-973X  
入藏号:  
CSCD:7275018  
Indexed Date:  
2022-11-07  
Institution Address:  
火箭军工程大学导弹工程学院, 西安, 陕西 710025, 中国  
Institution Address:  
College of Missile Engineering,Rocket Force University of Engineering, Xi'an, Shaanxi 710025, China  
Institution:  
火箭军工程大学导弹工程学院  
Institution:  
College of Missile Engineering,Rocket Force University of Engineering

### 466. LQCANet: Learnable-Query-Guided Multi-Scale Fusion Network Based on

摘要: Millimeter-wave radar semantic segmentation has proven successful in autonomous driving environment perception tasks. However, relying solely on range-angle (RA) images excludes Doppler information, which is crucial for dynamic target recognition. Despite various fusion method proposals, they consistently encounter issues with interfering RA information and poor fusion results. We introduce a novel learnable-query-guided multi-scale fusion network (LQCANet) for radar semantic segmentation, leveraging learnable-query for effective multi-scale cross-attention fusion. The cross-attention fusion module (CAF) initializes queries randomly, interacting with range-Doppler (RD) and angle-Doppler (AD) information via multi-layer cross-attention. Subsequently, the original RA features integrate with the generated queries to achieve multi-scale feature fusion. This approach prevents interference with RA information and ensures efficient fusion. Additionally, for enhanced feature extraction capabilities, this study introduces the pointwise multi-head self-attention down module (PMD), integrating a convolutional neural network (CNN) and a Transformer to extract both local and global features. Furthermore, pointwise convolution serves as an implicit positional coding method, addressing the limitation that explicit positional coding is not applicable to millimeter-wave radar images. Experiments demonstrate the superior performance of the proposed LQCANet on the CARRADA dataset. In comparison to the state-of-the-art (SOTA) fusion-network TMVA-Net, the LQCANet exhibits a significant improvement of 2.7 in mean intersection over union (mIoU) and 3.4 in mean dice similarity coefficient (mDice), while maintaining a computational complexity of only 27% (27.6 GFLOPs) of TMVA-Net. LQCANet achieves a superior trade-off between detection accuracy and speed, rendering it more suitable for environment perception tasks.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
2379-8858  
eISSN:  
2379-8904  
入藏号:  
WOS:001215322100016  
Indexed Date:  
2024-06-27  
Institution Address:  
Anhui Univ, Sch Integrated Circuits, Hefei 230039, Peoples R China  
Institution Address:  
Changshu Inst Technol, Sch Elect Informat Engn, Changshu 215505, Peoples R China  
Institution:  
Institution:

### 467. Combining convolutional neural networks and self-attention for fundus

摘要: Early detection of lesions is of great significance for treating fundus diseases. Fundus photography is an effective and convenient screening technique by which common fundus diseases can be detected. In this study, we use color fundus images to distinguish among multiple fundus diseases. Existing research on fundus disease classification has achieved some success through deep learning techniques, but there is still much room for improvement in model evaluation metrics using only deep convolutional neural network (CNN) architectures with limited global modeling ability; the simultaneous diagnosis of multiple fundus diseases still faces great challenges. Therefore, given that the self-attention (SA) model with a global receptive field may have robust global-level feature modeling ability, we propose a multistage fundus image classification model MBSaNet which combines CNN and SA mechanism. The convolution block extracts the local information of the fundus image, and the SA module further captures the complex relationships between different spatial positions, thereby directly detecting one or more fundus diseases in retinal fundus image. In the initial stage of feature extraction, we propose a multiscale feature fusion stem, which uses convolutional kernels of different scales to extract low-level features of the input image and fuse them to improve recognition accuracy. The training and testing were performed based on the ODIR-5k dataset. The experimental results show that MBSaNet achieves state-of-the-art performance with fewer parameters. The wide range of diseases and different fundus image collection conditions confirmed the applicability of MBSaNet.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
19  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
21  
ISSN:  
2045-2322  
入藏号:  
WOS:000907164800045  
PubMed ID:  
36593268  
Indexed Date:  
2023-02-04  
Institution Address:  
Chongqing Univ Technol, Sch Artificial Intelligence, Chongqing 401135, Peoples R China  
Institution Address:  
Chongqing Normal Univ, Coll Comp & Informat Sci, Chongqing 401331, Peoples R China  
Institution:  
Institution:

### 468. A real-time and high-accuracy railway obstacle detection method using

摘要: With the sustained development of railway transportation, the urgent necessity to improve train operation safety makes obstacle detection become the research focus. However, existing railway obstacle detectors still face challenges in balancing detection accuracy and speed during the shunting process. In addition, they are not robust enough in real-world railway environments, especially in complex scenes involving small obstacles. To address these problems, this paper presents a real-time and high-accuracy railway obstacle detection model using lightweight CNN and improved transformer (RH-Net) for detecting railway obstacles efficiently to guarantee traffic safety. First, the Lightweight Feature Extraction Module (LEM) is designed to minimize the model's computational load while maintaining its feature extraction ability. Then, the Improved Transformer Module (IFM) is developed to boost the model's ability about stably extract global contextual information. Finally, the Enhanced Multi-Scale Feature Fusion Module (EFM) is proposed to optimize the detection of obstacles with different sizes, especially small objects. In the experiments on railway dataset, RH-Net achieves optimal detection performance on GeForce GTX 1080Ti (96.99% mAP and 135 FPS) and Jetson Xavier NX (97.02% mAP and 43 FPS), which is significantly superior to the existing detection models. Experimental results show that RH-Net has excellent detection ability, which can accurately and efficiently detect obstacles in complicated railway environments. Moreover, the experiments on MS COCO indicate that RH-Net can achieve more satisfactory detection performance than existing state-of-the-art methods. Therefore, the proposed model can be well-applied to more complex real-world scenes for multiple object detection.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
0263-2241  
eISSN:  
1873-412X  
入藏号:  
WOS:001280638400001  
Indexed Date:  
2024-08-07  
Institution Address:  
Tianjin Univ, State Key Lab Precis Measuring Technol & Instrumen, Tianjin 300072, Peoples R China  
Institution Address:  
China Univ Min & Technol, Sch Mech Elect & Informat Engn, Beijing 100083, Peoples R China  
Institution:  
Institution:

### 469. Multi-scale multi-attention network for diabetic retinopathy grading

摘要: Objective. Diabetic retinopathy (DR) grading plays an important role in clinical diagnosis. However, automatic grading of DR is challenging due to the presence of intra-class variation and small lesions. On the one hand, deep features learned by convolutional neural networks often lose valid information about these small lesions. On the other hand, the great variability of lesion features, including differences in type and quantity, can exhibit considerable divergence even among fundus images of the same grade. To address these issues, we propose a novel multi-scale multi-attention network (MMNet). Approach. Firstly, to focus on different lesion features of fundus images, we propose a lesion attention module, which aims to encode multiple different lesion attention feature maps by combining channel attention and spatial attention, thus extracting global feature information and preserving diverse lesion features. Secondly, we propose a multi-scale feature fusion module to learn more feature information for small lesion regions, which combines complementary relationships between different convolutional layers to capture more detailed feature information. Furthermore, we introduce a Cross-layer Consistency Constraint Loss to overcome semantic differences between multi-scale features. Main results. The proposed MMNet obtains a high accuracy of 86.4% and a high kappa score of 88.4% for multi-class DR grading tasks on the EyePACS dataset, while 98.6% AUC, 95.3% accuracy, 92.7% recall, 95.0% precision, and 93.3% F1-score for referral and non-referral classification on the Messidor-1 dataset. Extensive experiments on two challenging benchmarks demonstrate that our MMNet achieves significant improvements and outperforms other state-of-the-art DR grading methods. Significance. MMNet has improved the diagnostic efficiency and accuracy of diabetes retinopathy and promoted the application of computer-aided medical diagnosis in DR screening.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0031-9155  
eISSN:  
1361-6560  
入藏号:  
WOS:001129432600001  
PubMed ID:  
38035368  
Indexed Date:  
2024-01-04  
Institution Address:  
Guangxi Normal Univ, Sch Elect & Informat Engn, Guilin 541004, Peoples R China  
Institution Address:  
Guangxi Normal Univ, Sch Comp Sci & Engn, Guilin 541004, Peoples R China  
Institution:  
Institution:

### 470. Joint learning strategy of multi-scale multi-task convolutional neural

摘要: Remaining useful life (RUL) prediction and health status (HS) assessment are two key tasks in aero-engine prognostics and health management (PHM) system. However, existing deep learning-based prognostic models perform RUL prediction and HS assessment tasks separately, without considering the correlation between these two tasks. Secondly, traditional deep learning can only extract single-scale features, which limits the ability to extract complex degradation features from high-dimensional condition monitoring data. Therefore, this work proposes a multi-scale and multi-task convolutional neural network for joint learning of aero-engine RUL prediction and HS assessment. Firstly, multi-sensor data with multiple cycles are converted into image samples to integrate more condition monitoring information that is beneficial to prognosis. Then, the multi-scale feature fusion block is designed as the shared network for multi-task, utilizing convolutional layers with filters of different sizes to enhance the ability to extract complex degradation features from high-dimensional condition monitoring data. And a multi-layer concatenation block is constructed to integrate multi-scale features at different levels to fully utilize the important information at different levels. On this basis, a multi-task joint learning block is constructed and a joint loss function is developed for joint learning of RUL prediction and HS assessment. Finally, experiments on two engine degradation datasets, CMAPSS and N-CMAPSS, demonstrate that the proposed network has excellent RUL prediction and HS assessment performance, and outperforms other state-of-the-art methods.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
9  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
ISSN:  
1568-4946  
eISSN:  
1872-9681  
入藏号:  
WOS:001241214600001  
Indexed Date:  
2024-06-13  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Civil Aviat, Nanjing 211106, Peoples R China  
Institution Address:  
Shandong Normal Univ, Sch Publ Adm, Jinan 250358, Peoples R China  
Institution:  
Institution:

### 471. A dual encoder LDCT image denoising model based on cross-scale skip

摘要: LDCT image denoising is crucial in medical imaging as it aims to minimize patient radiation exposure while maintaining diagnostic image quality. However, current convolutional neural network-based denoising methods struggle to incorporate global contexts, often focusing solely on local features. This limitation poses a significant challenge. To address this, a dual encoder denoising model is introduced that utilizes the Transformer model's proficiency in capturing long-range dependencies and global context. This model integrates the Transformer branch and the convolutional branch in the encoder. By concatenating the features of these two different branches, the model can capture both global and local image features, substantially enhancing denoising efficacy. A cross-scale skip connection mechanism is introduced to integrate the encoder's low-level features with the decoder' s high-level features, enriching contextual information and preserving image details. In addition, to meet the requirements of multi-scale feature fusion, the decoder is equipped with different multi-scale convolution modules to optimize feature processing. The number of layers in these modules gradually decreases as the depth of the decoder increases. In order to enhance the discriminative ability of the model, a multi-scale discriminator is also introduced, which effectively improves the recognition ability of the image by extracting features from four different scales. Consequently, our approach demonstrates remarkable performance in reducing noise and improving LDCT image quality, as evidenced by the substantial improvements in PSNR (17.75%) and SSIM (7.31%) values.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:001368551500001  
Indexed Date:  
2024-12-09  
Institution Address:  
Taiyuan Univ Sci & Technol, Coll Comp Sci & Technol, Taiyuan 030024, Shanxi, Peoples R China  
Institution Address:  
Taiyuan Univ Sci & Technol, Coll Elect Informat Engn, Taiyuan 030024, Shanxi, Peoples R China  
Institution:  
Institution:

### 472. Computer-aided diagnosis system based on multi-scale feature fusion for

摘要: For endoscopists, large-scale screening of gastrointestinal (GI) diseases is arduous and time-consuming. While their workload and human factor-induced errors can be reduced by computer-aided diagnosis (CAD) systems, the existing ones mainly focus on a limited number of lesions or specific organs, making them unsuitable for diagnosing various GI diseases in large-scale disease screening. This paper proposes a transformer and convolutional neural network-based CAD system (called TransMSF) to assist endoscopists in diagnosing multiple GI diseases. This system constructs two feature extraction paths with different coding methods to obtain the lesions' global and local information. In addition, downsampling is implemented in transformer to get global information of different scales, further enriching the feature representation and reducing the amount of computation and memory occupation. Moreover, a channel and spatial attention module with fewer parameters was successfully designed to pay more attention to the target and reduce the loss of important information during spatial dimension transformation. Finally, the extracted feature information is fused through the feature fusion module and then input into the linear classifier for disease diagnosis. The proposed system outperformed that of other state-of-the-art models on two datasets, reaching a 98.41% precision, a 98.15% recall, a 98.13% accuracy, and a 98.28% F1 score on the in-house GI dataset versus a 95.88% precision, a 95.88% recall, a 98.97% accuracy, and a 95.88% F1 score on the public Kvasir dataset. Moreover, TransMSF's performance was superior to that of seasoned endoscopists. The above results prove that the proposed system is instrumental in diagnosing GI diseases in large-scale disease screening. It can also be used as a training tool for junior endoscopists to improve their professional skills by rendering helpful suggestions.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2288-5048  
入藏号:  
WOS:000928230000002  
Indexed Date:  
2023-03-07  
Institution Address:  
Shandong Univ, Sch Control Sci & Engn, Jinan 250061, Peoples R China  
Institution Address:  
Shandong First Med Univ, Shandong Prov Hosp, Dept Gastroenterol, Jinan 250021, Peoples R China  
Institution:  
Institution:

### 473. BGRD-TransUNet: A Novel TransUNet-Based Model for Ultrasound Breast

摘要: Breast UltraSound (BUS) imaging is a commonly used diagnostic tool in the field of counter fighting breast diseases, especially for early detection and diagnosis of breast cancer. Due to the inherent characteristics of ultrasound images such as blurry boundaries and diverse tumor morphologies, it is challenging for doctors to manually segment breast tumors. In recent years, the Convolutional Neural Network (CNN) technology has been widely applied to automatically segment BUS images. However, due to the inherent limitations of CNNs in capturing global contextual information, it is difficult to capture the full context. To address this issue, the paper proposes a novel BGRD-TransUNet model for breast lesion segmentation, based on TransUNet. The proposed model, first, replaces the original ResNet50 backbone network of TransUNet with DenseNet121 for initial feature extraction. Next, newly designed Residual Multi-Scale Feature Modules (RMSFMs) are employed to extract features from various layers of DenseNet121, thus capturing richer features within specific layers. Thirdly, a Boundary Guidance (BG) network is added to enhance the contour information of BUS images. Additionally, newly designed Boundary Attentional Feature Fusion Modules (BAFFMs) are used to integrate edge information and features extracted through RMSFMs. Finally, newly designed Parallel Channel and Spatial Attention Modules (PCSAMs) are used to refine feature extraction using channel and spatial attention. An extensive experimental testing performed on two public datasets demonstrates that the proposed BGRD-TransUNet model outperforms all state-of-the-art medical image segmentation models, participating in the experiments, according to all evaluation metrics used (except for few separate cases), including the two most important and widely used metrics in the field of medical image segmentation, namely the Intersection over Union (IoU) and Dice Similarity Coefficient (DSC). More specifically, on the BUSI dataset and dataset B, BGRD-TransUNet achieves IoU values of 76.77% and 86.61%, and DSC values of 85.08% and 92.47%, respectively, which are higher by 7.27 and 3.64, and 5.81 and 2.54 percentage points, than the corresponding values achieved by the baseline (TransUNet).  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
2169-3536  
入藏号:  
WOS:001174988400001  
Indexed Date:  
2024-03-28  
Institution Address:  
North China Univ Sci & Technol, Hebei Key Lab Ind Intelligent Percept, Tangshan 063210, Peoples R China  
Institution Address:  
Univ Limerick, Telecommun Res Ctr TRC, Limerick V94 T9PX, Ireland  
Institution Address:  
Tangshan Univ, Intelligence & Informat Engn Coll, Tangshan 063000, Peoples R China  
Institution Address:  
Xian Jiaotong Liverpool Univ, Dept Comp, Suzhou 215000, Peoples R China  
Institution Address:  
Shenzhen Univ, Sch Biomed Engn, Hlth Sci Ctr, Shenzhen 518060, Guangdong, Peoples R China  
Institution Address:  
Plovdiv Univ Paisii Hilendarski, Dept Comp Syst, Plovdiv 4000, Bulgaria  
Institution Address:  
Bulgarian Acad Sci, Inst Math & Informat, Sofia 1040, Bulgaria  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 474. EIEN: Endoscopic Image Enhancement Network Based on Retinex Theory

摘要: In recent years, deep convolutional neural network (CNN)-based image enhancement has shown outstanding performance. However, due to the problems of uneven illumination and low contrast existing in endoscopic images, the implementation of medical endoscopic image enhancement using CNN is still an exploratory and challenging task. An endoscopic image enhancement network (EIEN) based on the Retinex theory is proposed in this paper to solve these problems. The structure consists of three parts: decomposition network, illumination correction network, and reflection component enhancement algorithm. First, the decomposition network model of pre-trained Retinex-Net is retrained on the endoscopic image dataset, and then the images are decomposed into illumination and reflection components by this decomposition network. Second, the illumination components are corrected by the proposed self-attention guided multi-scale pyramid structure. The pyramid structure is used to capture the multi-scale information of the image. The self-attention mechanism is based on the imaging nature of the endoscopic image, and the inverse image of the illumination component is fused with the features of the green and blue channels of the image to be enhanced to generate a weight map that reassigns weights to the spatial dimension of the feature map, to avoid the loss of details in the process of multi-scale feature fusion and image reconstruction by the network. The reflection component enhancement is achieved by sub-channel stretching and weighted fusion, which is used to enhance the vascular information and image contrast. Finally, the enhanced illumination and reflection components are multiplied to obtain the reconstructed image. We compare the results of the proposed method with six other methods on a test set. The experimental results show that EIEN enhances the brightness and contrast of endoscopic images and highlights vascular and tissue information. At the same time, the method in this paper obtained the best results in terms of visual perception and objective evaluation.  
Web of Science 核心合集中的 "被引频次":  
7  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
9  
eISSN:  
1424-8220  
入藏号:  
WOS:000831900900001  
PubMed ID:  
35891145  
Indexed Date:  
2022-08-05  
Institution Address:  
Anhui Univ, Sch Integrated Circuits, Hefei 230601, Peoples R China  
Institution Address:  
AnHui Engn Lab Agroecol Big Data, Hefei 230601, Peoples R China  
Institution:  
Institution:  
AnHui Engn Lab Agroecol Big Data

### 475. YOLOv8-GABNet: An Enhanced Lightweight Network for the High-Precision

摘要: Existing deep learning models for detecting citrus diseases and nutritional deficiencies grapple with issues related to recognition accuracy, complex backgrounds, occlusions, and the need for lightweight architecture. In response, we developed an improved YOLOv8-GABNet model designed specifically for citrus disease and nutritional deficiency detection, which effectively addresses these challenges. This model incorporates several key enhancements: A lightweight ADown subsampled convolutional block is utilized to reduce both the model's parameter count and its computational demands, replacing the traditional convolutional module. Additionally, a weighted Bidirectional Feature Pyramid Network (BiFPN) supersedes the original feature fusion network, enhancing the model's ability to manage complex backgrounds and achieve multiscale feature extraction and integration. Furthermore, we introduced important features through the Global to Local Spatial Aggregation module (GLSA), focusing on crucial image details to enhance both the accuracy and robustness of the model. This study processed the collected images, resulting in a dataset of 1102 images. Using LabelImg, bounding boxes were applied to annotate leaves affected by diseases. The dataset was constructed to include three types of citrus diseases-anthracnose, canker, and yellow vein disease-as well as two types of nutritional deficiencies, namely magnesium deficiency and manganese deficiency. This dataset was expanded to 9918 images through data augmentation and was used for experimental validation. The results show that, compared to the original YOLOv8, our YOLOv8-GABNet model reduces the parameter count by 43.6% and increases the mean Average Precision (mAP50) by 4.3%. Moreover, the model size was reduced from 50.1 MB to 30.2 MB, facilitating deployment on mobile devices. When compared with mainstream models like YOLOv5s, Faster R-CNN, SSD, YOLOv9t, and YOLOv10n, the YOLOv8-GABNet model demonstrates superior performance in terms of size and accuracy, offering an optimal balance between performance, size, and speed. This study confirms that the model effectively identifies the common diseases and nutritional deficiencies of citrus from Conghua's "Citrus Planet". Future deployment to mobile devices will provide farmers with instant and precise support.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2077-0472  
入藏号:  
WOS:001363596500001  
Indexed Date:  
2024-12-01  
Institution Address:  
South China Agr Univ, Coll Elect Engn, Coll Artificial Intelligence, Guangzhou 510642, Peoples R China  
Institution Address:  
China Agr Res Syst, Div Citrus Machinery, Guangzhou 510642, Peoples R China  
Institution Address:  
Guangdong Engn Res Ctr Monitoring Agr Informat, Guangzhou 510642, Peoples R China  
Institution:  
Institution:  
China Agr Res Syst  
Institution:  
Guangdong Engn Res Ctr Monitoring Agr Informat

### 476. Fault Diagnosis Method of Rolling Bearing Based on MSCNN-LSTM

摘要: Deep neural networks have been widely applied to bearing fault diagnosis systems and achieved impressive success recently. To address the problem that the insufficient fault feature extraction ability of traditional fault diagnosis methods results in poor diagnosis effect under variable load and noise interference scenarios, a rolling bearing fault diagnosis model combining Multi-Scale Convolutional Neural Network (MSCNN) and Long Short-Term Memory (LSTM) fused with attention mechanism is proposed. To adaptively extract the essential spatial feature information of various sizes, the model creates a multi-scale feature extraction module using the convolutional neural network (CNN) learning process. The learning capacity of LSTM for time information sequence is then used to extract the vibration signal's temporal feature information. Two parallel large and small convolutional kernels teach the system spatial local features. LSTM gathers temporal global features to thoroughly and painstakingly mine the vibration signal's characteristics, thus enhancing model generalization. Lastly, bearing fault diagnosis is accomplished by using the SoftMax classifier. The experiment outcomes demonstrate that the model can derive fault properties entirely from the initial vibration signal. It can retain good diagnostic accuracy under variable load and noise interference and has strong generalization compared to other fault diagnosis models.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001268119400015  
Indexed Date:  
2024-07-22  
Institution Address:  
Northeast Elect Power Univ, Key Lab Modern Power Syst Simulat & Control & Rene, Minist Educ, Jilin 132012, Peoples R China  
Institution Address:  
Northeast Elect Power Univ, Minist Educ, Jilin 132012, Peoples R China  
Institution:  
Institution:

### 477. Multimodal Emotion Recognition Using a Hierarchical Fusion Convolutional

摘要: In recent years, deep learning has been increasingly used in the field of multimodal emotion recognition in conjunction with electroencephalogram. Considering the complexity of recording electroencephalogram signals, some researchers have applied deep learning to find new features for emotion recognition. In previous studies, convolutional neural network model was used to automatically extract features and complete emotion recognition, and certain results were obtained. However, the extraction of hierarchical features with convolutional neural network for multimodal emotion recognition remains unexplored. Therefore, this paper proposes a hierarchical fusion convolutional neural network model to mine the potential information in the data by constructing different network hierarchical structures, extracting multiscale features, and using feature-level fusion to fuse the global features formed by combining weights with manually extracted statistical features to form the final feature vector. This paper conducts binary classification experiments on the valence and arousal dimensions of the DEAP and MAHNOB-HCI data sets to evaluate the performance of the proposed model. The results show that the model proposed in this paper can achieve accuracies of 84.71% and 89.00% on the two corresponding data sets, indicating that the model proposed in this paper is superior to other deep learning emotion classification models in feature extraction and fusion.  
Web of Science 核心合集中的 "被引频次":  
39  
在中国科学引文数据库中的被引频次:  
3  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
42  
ISSN:  
2169-3536  
入藏号:  
WOS:000608193200001  
Indexed Date:  
2021-01-29  
Institution Address:  
Liaoning Normal Univ, Sch Comp & Informat Technol, Dalian 116081, Peoples R China  
Institution:

### 478. SpAtNet: a spatial feature attention network for hand gesture

摘要: Hand Gesture Recognition is receiving enormous attention because they facilitate communication for various applications, including human-computer interaction. However, the HGR system poses various challenges due to environmental conditions, rotation, scaling, illumination variations, etc. This paper proposes a lightweight CNN based portable network SpAtNet: a spatial feature attention network that learns spatial features for precise hand gesture recognition. SpAtNet primarily consists of two blocks: multi-scale attentive feature fusion (MAFF) and interleaved module. The MAFF block employs multi-scale filters: 1x1, 3x3, 5x5 to extract the rich spatial information, which improves the robustness of the HGR system. The MAFF block encodes features with smaller scale utilizing small filters while a larger filter extracts coarse features. The interleaved module is designed by sequentially stacking four convolutional layers with kernel sizes: 3x3 and 5x5. The interleaved module is introduced to learn the high-level contextual features crucial for efficient recognition. The proposed algorithm is validated on six benchmark datasets: MUGD, ASL Finger Spelling, NUS-II, HGR-I, Triesch and ArASL. The comparative analysis and visual representation show that the proposed approach outperforms the other state-of-art techniques.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
22  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
在中国科学引文数据库中的被引频次:  
1  
被引频次合计:  
23  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001142522500005  
Indexed Date:  
2024-03-02  
Institution Address:  
Natl Inst Technol Jamshedpur, CSE Dept, Jamshedpur, Jharkhand, India  
Institution Address:  
Natl Inst Technol Sikkim, CSE Dept, Ravangla, Sikkim, India  
Institution:  
Institution:

### 479. An improved SSD lightweight network with coordinate attention for

摘要: Accurate identification and monitoring of aircraft on the airport surface can assist managers in rational scheduling and reduce the probability of aircraft conflicts, an important application value for constructing a "smart airport." For the airport surface video monitoring, there are small aircraft targets, aircraft obscuring each other, and affected by different weather, the aircraft target clarity is low, and other complex monitoring problems. In this paper, a lightweight model network for video aircraft recognition in airport field video in complex environments is proposed based on SSD network incorporating coordinate attention mechanism. First, the model designs a lightweight feature extraction network with five feature extraction layers. Each feature extraction layer consists of two modules, Block A and Block I. The Block A module incorporates the coordinate attention mechanism and the channel attention mechanism to improve the detection of obscured aircraft and to enhance the detection of small targets. The Block I module uses multi-scale feature fusion to extract feature information with rich semantic meaning to enhance the feature extraction capability of the network in complex environments. Then, the designed feature extraction network is applied to the improved SSD detection algorithm, which enhances the recognition accuracy of airport field aircraft in complex environments. It was tested and subjected to ablation experiments under different complex weather conditions. The results showthat compared with the Faster R-CNN, SSD, andYOLOv3models, the detection accuracy of the improved model has been increased by 3.2%, 14.3%, and 10.9%, respectively, and the model parameters have been reduced by 83.9%, 73.1%, and 78.2% respectively. Compared with the YOLOv5 model, the model parameters are reduced by 38.9% when the detection accuracy is close, and the detection speed is increased by 24.4%, reaching 38.2fps, which can well meet the demand for real-time detection of aircraft on airport surfaces.  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1064-1246  
eISSN:  
1875-8967  
入藏号:  
WOS:001163267400025  
Indexed Date:  
2024-04-06  
Institution Address:  
Henan Univ Technol, Minist Educ, Key Lab Grain Informat Proc & Control, Zhengzhou, Peoples R China  
Institution Address:  
Henan Univ Technol, Henan Key Lab Grain Photoelect Detect & Control, Zhengzhou, Peoples R China  
Institution Address:  
Shenzhen Univ, Key Lab Geo Environm Monitoring Great Bay Area, MNR, Shenzhen, Peoples R China  
Institution:  
Institution:  
Institution:

### 480. Design of a Lung Lesion Target Detection Algorithm Based on a

摘要: This study developed a novel domain-adaptive neural network framework, CNDAD-Net, for addressing the challenges of lung lesion detection in cross-domain medical image analysis. The proposed framework integrates domain adaptation techniques into a classical encoding-decoding structure to align feature distributions between source and target domains. Specifically, a "Generative Adversarial Network" GAN-based domain discriminator is utilized for the iterative refinement of feature representations to minimize cross-domain discrepancies and improve the generalization capability of the model. In addition, a novel Cross-Fusion Block (CFB) is proposed to implement multi-scale feature fusion that facilitates the deep integration of 2D, 3D, and domain-adapted features. The CFB achieves bidirectional feature flow across dimensions, thereby improving the model's capability to detect diverse lesion morphologies while minimizing false positives and missed detections. For better detection, coarse-grained domain adaptation is implemented by MMD for further optimization. It integrates a module inspired by a CycleGAN for the process to generate high-resolution images on low-quality data. Using the Lung Nodule Analysis (LUNA16) dataset, the test was conducted and its experimental result was compared with that of previous standard methods such as Faster R-CNN and YOLO, yielding mAP 0.889, recall at 0.845 and the F1-score at 0.886. This work, with a novel CNDAD-Net model, lays down a solid and scalable framework for the precise detection of lung lesions, which is extremely critical for early diagnosis and treatment. The model has prospects and is capable of being extended in future to multimodal imaging data ad real-time diagnostic scenarios, and can help in further developing intelligent medical image analysis systems.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001442392400001  
Indexed Date:  
2025-03-18  
Institution Address:  
City Univ Macau, Fac Data Sci, Macau 999078, Peoples R China  
Institution Address:  
Macao Polytech Univ, Fac Appl Sci, Macau 999078, Peoples R China  
Institution:  
Institution:

### 481. Deep learning-based borescope image processing for aero-engine blade

摘要: As an important part of aero-engines, blades are used to compress the air entering the engine and produce a lot of thrusts. Catastrophic and sudden accidents caused by blade failure seriously threaten aero-engine operation safety, hence it is necessary to check blades regularly to ensure their health and reliability. As a common non-destructive testing technology, borescope inspection is widely used in health monitoring and maintenance of aero-engine blades. However, traditional borescope inspection mainly relies on artificial vision and it is a time-consuming and experience-dependent process. In this paper, a deep learning-based blade damage detection method is proposed to endow borescope inspection with intelligence. The proposed method pays more attention to texture information, which reflects the types of damages. It is applicable to the situation requiring higher localization accuracy due to the balance between coarse-grained and fine-grained localization. In this method, the enhanced Mask R CNN network with three functions of damage mode classification, damage localization, and damage area segmentation is constructed. Moreover, a texture-focus multi-scale feature fusion network is used to give more attention to the shallow texture information which reflects the shape of damage. Balanced L1 loss is introduced to balance coarse-grained and fine-grained localization by adjusting the gradient and loss of easy samples. We also propose practical evaluation metrics for blade damage detection and make detailed evaluations and discussions. Extensive experiments are conducted on simulated and real aeroengine damaged blade datasets to verify the effectiveness and progressiveness of our method, and the results show the method has great potential for intelligent detection of aero-engine in-situ blade damage.(c) 2022 Elsevier Masson SAS. All rights reserved.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
5  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
50  
被引频次合计:  
55  
ISSN:  
1270-9638  
eISSN:  
1626-3219  
入藏号:  
WOS:000793408800008  
Indexed Date:  
2022-05-19  
Institution Address:  
Xi An Jiao Tong Univ, Sch Mech Engn, Xian 710049, Shaanxi, Peoples R China  
Institution:

### 482. Dental model segmentation network with fine-grained receptive fields and

摘要: Objective Dental computer-aided therapy relies on the use of dental models to aid dentists in their practice.One of the most fundamental tasks in dental computer-aided therapy is the automated segmentation of teeth using point cloud data obtained from intra-oral scanners(IOS).The precise segmentation of each individual tooth in this procedure provides vital information for a variety of subsequent tasks.These segmented dental models facilitate customized treatment planning and modeling,thus providing extensive assistance in carrying out further treatments.However,the automated segmentation of individual teeth from dental models faces three significant challenges.First,the indistinct boundary between teeth and gums poses difficulties in segmentation based solely on geometric features.Second,certain factors,such as occlusion during scanning,can lead to suboptimal results,particularly in posterior dental regions,thereby further complicating the segmentation process.Lastly,teeth often exhibit complex anomalies in patients,including crowding,missing teeth,and misalignment issues,which further complicate the task of accurate segmentation.To address these challenges,two conventional methods are proposed for segmenting teeth in images obtained from IOS scanners.The first method employs a projection-based approach,wherein a 3D dental scan image is initially projected into a 2D space,segmentation is then performed in a 2D space,and the result is remapped back into the 3D space.The second method adopts a geometrybased approach and typically utilizes geometric attributes,such as surface curvature,geodesic information,harmonic fields,and other geometric properties,to distinguish tooth structures.However,these methods are not fully automated and rely on domain-specific knowledge and experience.Moreover,the predefined low-level attributes used by these methods lack robustness when dealing with the complex appearance of patietns' teeth.Considering the impactful application of convolutional neural networks (CNN) in computer vision and medical image processing,several deep learning methods rooted in CNN have been introduced.Some of these methods directly extract translation-invariant depth geometric features from 3D point cloud data but suffer from a lack of necessary receptive field for fine-grained tasks,such as dental model segmentation.Moreover,the network structure exhibits redundancy and neglects the crucial details of dental models.To address these issues,a fully automatic tooth segmentation network called TRNet is proposed in this paper,which can automatically segment teeth on unprocessed intra-oral scanned point cloud models.Method In the proposed end-to-end 3D point cloudbased multi-scale fusion dental model segmentation method,an encoder with a fine-grained receptive field is employed to address those challenges posed by the small size of each tooth within the dental model and the lack of distinct features between the teeth and gums.Each tooth within the dental model is relatively small in comparison to the entire dental model,and the boundaries between the teeth and gums lack distinct features.Consequently,a fine-grained receptive field is essential for extracting features from this model.The network adopts a small radius for querying the neighborhood,thus narrowing the receptive field and enabling the network to focus on detailed features.Additionally,downsampling can lead to the uneven density of the point cloud,thereby causing the network trained on sparse point clouds to struggle in recognizing fine-grained local structures.Multiscale feature fusion coding is implemented to address these issues.Given that the encoder uses a small query radius to create a fine-grained receptive field,the relative coordinates become relatively small.Consequently,the network needs to learn large weights to operate on these relative coordinates,thereby introducing further challenges in network optimization.TRNet normalizes the relative coordinates in the feature extraction layer to facilitate network optimization and enhance segmentation performance.The network also employs a highly efficient decoder.Previous segmentation methods often utilize the U-Net structure,which incorporates jump connections for multi-level feature aggregation between the input features of the cascaded decoder and the outputs of the corresponding layer encoder.However,this top-down propagation is considered inefficient for feature aggregation.The decoding approach used by TRNet directly combines the features outputted from all cascade encoders,thereby allowing the network to learn the importance of each cascade.The discrepancies in scales or dimensions of the features represented by fused information in the network may also introduce unwanted bias during the fusion process.To address these issues and ensure that the network focuses on crucial information within the fused features,a soft attention mechanism is incorporated into the fusion process.Specifically,a soft attention operation is performed on the newly combined features after their connection,thereby enabling the network to adaptively balance the discrepancies of different scales or levels in the propagated features.Result A dataset comprising dental models taken from numerous patients with irregular tooth shapes,such as crowding,misalignment,and underdeveloped teeth,was compiled.To establish the labeled values,an experienced dentist meticulously segmented and annotated these models.The dataset was then randomly divided into two subsets,with 146 models allocated for training and 20 models reserved for testing.Data augmentation techniques,such as random panning and scaling,were employed to enhance the diversity of the training set.In each iteration,intra-oral scan images were shifted by a randomly selected value within the range of [-0.1,0.1] and scaled by a randomly chosen magnification within the range of [0.8,1.25],thereby generating new training data.Experimental results from a 5-fold cross-validation reveal that TRNet achieved an overall accuracy (OA) of 97.0150.096% and a mean intersection over union (mIoU) of 92.6910.454%,significantly outperforming the existing methods.Conclusion An end-to-end deep learning network called TRNet is introduced in this paper for the automatic segmentation of teeth in 3D dental images acquired from intra-oral scanners.An encoder with fine-grained receptive fields was also implemented to enhance the local feature extraction capabilities essential for dental model segmentation.Additionally,a decoder based on hierarchical connections was employed to allow the network to decode efficiently by learning the significance of each level.This refinement significantly improves the precision of dental model segmentation.A soft attention mechanism was also integrated into the feature fusion process to enable the network to focus on key information within dental model features.Experimental results indicate that TRNet shows excellent performance on intraoral scanned point cloud models and enhances the ability of the network to segment dental models,thereby improving the accuracy of point cloud segmentation results.  
摘要:  
目的从口内扫描点云模型上精确分割牙齿是计算机辅助牙科治疗中重要的任务,但存在手动执行耗时且烦琐的问题。近年来,计算机视觉领域涌现出一些端到端的方法实现三维形状分割。然而,大多数方法没有注意到口腔分割需要网络具有更加细粒度的感受野,因此分割精度仍然受到限制。为了解决该问题,设计了一个端到端的具有细粒度感受野的全自动牙齿分割网络--TRNet,用于在未加工的口内扫描点云模型上自动分割牙齿。方法首先,TRNet使用了具有细粒度感受野的编码器,其基于多尺度融合从不同的尺度提取到更全面的口腔模型特征,并通过更适合口腔模型分割的细粒度分组查询半径以及具有相对坐标归一化的特征提取层来提升分割性能。其次,TRNet采用了基于层级连接的特征嵌入方式,网络学习到口腔模型中由各个局部区域到覆盖更大范围空间的关键特征,特征提取更全面,提升了网络的分割精度。同时,TRNet使用了基于软性注意力机制的特征融合方式,使网络更好地从融合特征中关注到口腔模型的关键信息。结果使用由口内扫描仪获取的患者口内扫描点云模型数据集评估了TRNet。经过5折交叉验证的实验结果中,TRNet的总体准确率(overall accuracy,OA)达到了97.0150.096%,平均交并比(mean intersection over union,mIoU)达到了92.6910.454%,显著优于现有方法。结论实验结果表明,提出的具有细粒度感受野的多尺度融合口腔分割模型在口内扫描点云模型上取得了较好表现,提高了网络对于口腔模型的分割能力,使点云分割结果更准确。  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1006-8961  
入藏号:  
CSCD:7871682  
Indexed Date:  
2025-03-08  
Institution Address:  
常州大学计算机与人工智能学院;;苏州大学附属第三医院\常州市第一人民医院, ;;, 常州;;常州, ;; 213164;;213003  
Institution Address:  
常州大学计算机与人工智能学院, 常州, 213164  
Institution Address:  
苏州大学附属第三医院\常州市第一人民医院, 常州, 213003  
Institution Address:  
苏州大学附属第三医院\常州市第一人民医院;;常州大学医学与健康工程学院, ;;, 常州;;常州, ;; 213003;;213164  
Institution Address:  
School of Computer Science and Artificial Intelligence,Changzhou University;;The Third Affiliated Hospital of Soochow University,The First People's Hospital of Changzhou, ;;, Changzhou;;Changzhou, ;; 213164;;213003  
Institution Address:  
School of Computer Science and Artificial Intelligence,Changzhou University, Changzhou, 213164  
Institution Address:  
The Third Affiliated Hospital of Soochow University,The First People's Hospital of Changzhou, Changzhou, 213003  
Institution Address:  
The Third Affiliated Hospital of Soochow University,The First People's Hospital of Changzhou;;School of Medical and Health Engineering,Changzhou University, ;;, Changzhou;;Changzhou, ;; 213003;;213164  
Institution:  
常州大学计算机与人工智能学院;;苏州大学附属第三医院\常州市第一人民医院  
Institution:  
常州大学计算机与人工智能学院  
Institution:  
苏州大学附属第三医院\常州市第一人民医院  
Institution:  
苏州大学附属第三医院\常州市第一人民医院;;常州大学医学与健康工程学院  
Institution:  
School of Computer Science and Artificial Intelligence,Changzhou University;;The Third Affiliated Hospital of Soochow University,The First People's Hospital of Changzhou  
Institution:  
School of Computer Science and Artificial Intelligence,Changzhou University  
Institution:  
The Third Affiliated Hospital of Soochow University,The First People's Hospital of Changzhou  
Institution:  
The Third Affiliated Hospital of Soochow University,The First People's Hospital of Changzhou;;School of Medical and Health Engineering,Changzhou University

### 483. Fan Blade Crack Detection Algorithm Based on Multi-Scale Feature Fusion

摘要: With the rapid development of social economy, energy consumption is growing tremendously so green energy such as wind energy has become widely used, thus promoting the construction of wind turbines. Due to the long-term use of the electro-mechanical unit, the traditional maintenance cost is too high. In order to quickly and accurately detect and maintain the fan blades, based on the intelligent big data from the environment, we propose the convolutional neural network model to solve the problem of low recognition rate due to the lack of feature extraction in the fan blade crack image, and the long short-term memory network (Long Short-Term Memory, LSTM) convolutional neural network model, and the dimensionality reduction of the captured image data, which is beneficial to improve the recognition rate of the picture and reduce the loss rate of the picture through the detection model's suitable recognition of complex background problems such as target occlusion and overlap. Using LSTM to extract the global context module can effectively improve the target detection accuracy. When this part is added, the detection accuracy will increase by about 3% to 7%. The image position can be accurately captured and the recognition rate is greatly improved through the optimized convolutional neural network, which can provide a reference for future research in other fields.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001410255600009  
Indexed Date:  
2025-02-09  
Institution Address:  
Guangdong Baiyun Univ, Fac Megadata & Comp, Guangzhou 510450, Peoples R China  
Institution Address:  
Mongolian Univ Sci & Technol, Sch Informat & Commun Technol, Ulaanbaatar 13341, Bayanzurkh, Mongolia  
Institution:  
Guangdong Baiyun Univ  
Institution:

### 484. CMFE-PVT: a lightweight fault diagnosis framework for rolling bearings

摘要: The intelligent fault diagnosis method based on transformer and convolutional neural network (CNN) has achieved good global and local feature extraction results. However, the multi-head self-attention mechanism adopted by the transformer and the cross-channel convolution operation in CNN increases the complexity of the model, thereby increasing the demand for hardware resources, which to some extent, limits its broad applicability in industrial applications. Therefore, this paper proposes a lightweight fault diagnosis framework based on compact multi-scale feature extraction and pruned-restructured vision transformer (ViT) to address the above challenges. Firstly, a compact multi-scale feature extraction module is designed to efficiently capture complex features in rolling bearing vibration signals through parallel multi-scale convolution kernels, combined with channel reduction strategies to significantly reduce computational complexity while maintaining feature richness. Next, short-time Fourier transform and pseudo-color processing techniques are used to obtain time-frequency images. Then, a dual optimization of matrix sparsity and structural reorganization is implemented for Self-attention in ViT to ensure model performance and significantly reduce computational overhead. Finally, the time-frequency images are segmented and rearranged before being fed into the improved lightweight ViT for global feature extraction and fault recognition of rolling bearings. The experimental results show that the proposed fault diagnosis method has the advantages of lightweight (Params:4.27 K, floating point operations per seconds:0.1 M, multiplication and accumulation operations per seconds:51.07 K) and robustness compared to mainstream algorithms.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001395522400001  
Indexed Date:  
2025-01-19  
Institution Address:  
Tongling Univ, Sch Elect Engn, Tongling, Anhui, Peoples R China  
Institution Address:  
Anhui Univ Technol, Coll Elect & Informat Engn, Maanshan, Anhui, Peoples R China  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Aerosp Engn, Nanjing, Jiangsu, Peoples R China  
Institution:  
Institution:  
Institution:

### 485. High-Resolution Remote Sensing Farmland Extraction Network Based on

摘要: Deep learning-based high-resolution remote sensing for farmland extraction is a crucial method for obtaining large-scale farmland information. However, variations in crop types, growth conditions, and factors such as narrow edges in farmland lead to lower extraction accuracy and inaccurate boundaries in high-resolution remote sensing. Therefore, this letter proposes a multibranch convolutional neural network (FFENet) that employs a dense-feature overlay fusion module (FFM) and an information homogeneity enhancement module. This network facilitates rapid extraction and dense fusion of information at various scales through the implementation of the dense FM, thereby enhancing the model's representation of global consistency and local features. The information homogeneity enhancement module further strengthens the information exchange between the bottom and top layers, improves the fusion of feature information across branches, and ensures consistent representation of internal farmland features while enhancing differentiation at the edges. The experimental results demonstrate that the proposed method effectively considers both internal global consistency and local variations in edge information, thereby ensuring the integrity of farmland plots and the continuity of the farmland edges. The quantitative evaluation of the dataset shows that the model performs well in farmland extraction, with overall accuracy (OA) and intersection over union (IoU) reaching 95.41% and 93.74% on the GF-2 dataset and 94.75% and 88.28% on the JL-1 dataset.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001425523600003  
Indexed Date:  
2025-02-24  
Institution Address:  
Anhui Univ, Inst Phys Sci & Informat Technol, Hefei 230031, Peoples R China  
Institution Address:  
Anhui Univ, Sch Resources & Environm Engn, Hefei 230601, Peoples R China  
Institution Address:  
Anhui Univ, Sch Artificial Intelligence, Hefei 230601, Peoples R China  
Institution Address:  
Anhui & Huaihe River Inst Hydraul Res, Hefei 230088, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Anhui & Huaihe River Inst Hydraul Res

### 486. Unsupervised Difference-Guided Adaptive Fusion Network for Change

摘要: It is challenging to accurately extract change features in polarimetric synthetic aperture radar (PolSAR) image change detection. Recently, many existing convolutional neural network (CNN)-based change detection algorithms show satisfactory performance, but error detection still exists in the acquired change features. In this article, we propose an unsupervised difference-guided adaptive fusion network (DGAFN), which contains three key modules, i.e., adaptive difference extraction module (ADEM), difference feature enhancement module (DFEM), and adaptive weighted fusion module (AWFM). First, ADEM extracts initial multiscale difference features through nonlocal attention mechanisms guided by bitemporal feature differences, which helps the network to capture the correlation and difference between bitemporal features more accurately. Second, in order to further refine multiscale difference features and reduce pseudo-changes, DFEM uses class activation maps (CAMs) and contextual feature extraction modules (CFEMs) to perceive more polarization contextual information in a larger scale range. Finally, in order to improve the recognition ability of change targets with different sizes, AWFM adaptively fuses enhanced multiscale difference features in spatial and channel dimensions by introducing attention mechanisms to obtain change features with multiscale complementary information. Comparative experimental results on five measured PolSAR datasets obtained by GaoFen-3 show that DGAFN is superior to other relevant advanced algorithms in change detection performance.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001246585400019  
Indexed Date:  
2024-06-27  
Institution Address:  
Xidian Univ, Sch Elect Engn, Remote Sensing Image Proc & Fus Grp, Xian 710071, Peoples R China  
Institution Address:  
Xidian Univ, Natl Key Lab Radar Signal Proc, Xian 710071, Peoples R China  
Institution:  
Institution:

### 487. Multi-scale Transformer-CNN domain adaptation network for complex

摘要: Despite the breakthroughs in deep neural network-based fault diagnosis, the model mismatch problem owing to the changes in data distribution remains challenging. To fuse deep features for cross-mode feature modeling, a Transformer-convolutional neural network (TrCNN) based multi-scale distribution alignment network is proposed. In the source domain stage, a concatenated structure of Transformer and convolutional neural network (CNN) extracts deep diagnostic information by combining global and local approaches. In the transfer stage, alignment is performed on the complex features extracted from different CNN substructures at multiple scales. Multi-scale feature alignment allows aligning information from various aspects while maintaining the discriminability of the data. The effectiveness and feasibility of the proposed method were demonstrated through experiments conducted on the Tennessee-Eastman (TE) process and industrial three-phase flow (TFF) equipment. (c) 2023 Elsevier Ltd. All rights reserved.  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
20  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
21  
ISSN:  
0959-1524  
eISSN:  
1873-2771  
入藏号:  
WOS:001073825300001  
Indexed Date:  
2023-10-22  
Institution Address:  
Beijing Univ Chem Technol, Coll Informat Sci & Technol, Beijing 100029, Peoples R China  
Institution Address:  
Minist Educ China, Engn Res Ctr Intelligent PSE, Beijing 100029, Peoples R China  
Institution:  
Institution:  
Minist Educ China

### 488. A High Resolution Convolutional Neural Network with Squeeze and

摘要: Automatic modulation classification (AMC) technology is one of the cutting-edge technologies in cognitive radio communications. AMC based on deep learning has recently attracted much attention due to its superior performances in classification accuracy and robustness. In this paper, we propose a novel, high resolution and multi-scale feature fusion convolutional neural network model with a squeeze-excitation block, referred to as HRSENet, to classify different kinds of modulation signals. The proposed model establishes a parallel computing mechanism of multi-resolution feature maps through the multi-layer convolution operation, which effectively reduces the information loss caused by downsampling convolution. Moreover, through dense skip-connecting at the same resolution and up-sampling or down-sampling connection at different resolutions, the low resolution representation of the deep feature maps and the high resolution representation of the shallow feature maps are simultaneously extracted and fully integrated, which is benificial to mine signal multi-level features. Finally, the feature squeeze and excitation module embedded in the decoder is used to adjust the response weights between channels, further improving classification accuracy of proposed model. The proposed HRSENet significantly outperforms existing methods in terms of classification accuracy on the public dataset "Over the Air" in signal-to-noise (SNR) ranging from -2dB to 20dB. The classification accuracy in the proposed model achieves 85.36% and 97.30% at 4dB and 10dB, respectively, with the improvement by 9.71% and 5.82% compared to LWNet. Furthermore, the model also has a moderate computation complexity compared with several state-of-the-art methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1673-5447  
入藏号:  
WOS:001351913800017  
Indexed Date:  
2024-12-18  
Institution Address:  
Bejing Forestry Univ, Sch Informat Sci & Technol, Beijing 100083, Peoples R China  
Institution Address:  
Natl Forestry & Grassland Adm, Engn Res Ctr Forestry Oriented Intelligent Inform, Beijing 100083, Peoples R China  
Institution Address:  
La Trobe Univ, Dept Comp Sci & Informat Technol, Melbourne, Vic 3086, Australia  
Institution Address:  
Univ Sydney, Sch Elect & Informat Engn, Sydney, NSW 2006, Australia  
Institution:  
Bejing Forestry Univ  
Institution:  
Natl Forestry & Grassland Adm  
Institution:  
Institution:

### 489. Reconstruction of physical field characteristics of underwater vehicle

摘要: The characteristics of underwater vehicle wakes are intricately linked to multiple parameters, such as speed, depth, and environmental factors. Obtaining wake characteristic information for various operating conditions solely through numerical simulation methods can result in computational space explosion, rendering the time and computational resource costs prohibitive. This paper harnesses the robust image processing capabilities of convolutional neural networks and incorporates strategies such as attention mechanisms, dilated convolution techniques, and multi-scale feature fusion to design and construct a neural network architecture. Through a data-driven approach, it reconstructs multiple physical wake field characteristics resulting from underwater vehicle, including underwater velocity fields, surface divergence fields, surface kelvin wake, and surface thermal wakes. The study establishes a "black box" mapping between relevant parameters and the physical fields of wakes. The results demonstrate that the constructed network model achieves high accuracy in capturing both the macroscopic structures and pixel-level details of various physical fields. In comparison with the truth-values, the average normalized root mean square errors for the underwater velocity field, surface divergence field, surface kelvin wake, and surface thermal wakes are 6.10%, 3.40%, 8.21%, and 10.96%, respectively. The average structural similarity index values are 0.955, 0.966, 0.923, and 0.904, respectively. The predicted results closely match the truth-values for each physical field characteristic, effectively addressing the challenges of nonlinearity and multi-scale feature extraction in predicting complex flow fields, offering support for the rapid forecasting of multi-dimensional and multi-physical field characteristics of underwater vehicle wakes.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1070-6631  
eISSN:  
1089-7666  
入藏号:  
WOS:001359618100013  
Indexed Date:  
2024-11-30  
Institution Address:  
Naval Univ Engn, Inst Noise & Vibrat, Wuhan 430033, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Sch Marine Sci & Technol, Xian 710072, Shanxi, Peoples R China  
Institution:  
Institution:

### 490. MCPA: multi-scale cross perceptron attention network for 2D medical

摘要: The UNet architecture, based on convolutional neural networks (CNN), has demonstrated its remarkable performance in medical image analysis. However, it faces challenges in capturing long-range dependencies due to the limited receptive fields and inherent bias of convolutional operations. Recently, numerous transformer-based techniques have been incorporated into the UNet architecture to overcome this limitation by effectively capturing global feature correlations. However, the integration of the Transformer modules may result in the loss of local contextual information during the global feature fusion process. In this work, we propose a 2D medical image segmentation model called multi-scale cross perceptron attention network (MCPA). The MCPA consists of three main components: an encoder, a decoder, and a Cross Perceptron. The Cross Perceptron first captures the local correlations using multiple Multi-scale Cross Perceptron modules, facilitating the fusion of features across scales. The resulting multi-scale feature vectors are then spatially unfolded, concatenated, and fed through a Global Perceptron module to model global dependencies. Considering the high computational cost of using 3D neural network models, and the fact that many important clinical data can only be obtained in two dimensions, our MCPA focuses on 2D medical image segmentation. Furthermore, we introduce a progressive dual-branch structure (PDBS) to address the semantic segmentation of the image involving finer tissue structures. This structure gradually shifts the segmentation focus of MCPA network training from large-scale structural features to more sophisticated pixel-level features. We evaluate our proposed MCPA model on several publicly available medical image datasets from different tasks and devices, including the open large-scale dataset of CT (Synapse), MRI (ACDC), and widely used 2D medical imaging datasets captured by fundus camera (DRIVE, CHASE\_\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$\\_$$\end{document}DB1, HRF), and OCTA (ROSE). The experimental results show that our MCPA model achieves state-of-the-art performance.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2199-4536  
eISSN:  
2198-6053  
入藏号:  
WOS:001380505300008  
Indexed Date:  
2024-12-28  
Institution Address:  
Univ Sci & Technol China, Sch Biomed Engn, Div Life Sci & Med, Hefei, Peoples R China  
Institution Address:  
Univ Sci & Technol China, Suzhou Inst Adv Res, Suzhou, Peoples R China  
Institution Address:  
Univ Sci & Technol China, Dept Precis Machinery & Precis Instrument, Hefei, Peoples R China  
Institution Address:  
Univ Sci & Technol China, Sch Microelect, Hefei, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 491. Multi-attention embedded network for salient object detection

摘要: Although the salient object detection method based on the fully convolutional neural network has achieved better performance, how to learn effective feature representations in complex scenes to obtain more accurate saliency maps is still a challenge. In order to cope with the above-mentioned challenges, an addition or cascade structure is generally used to fuse feature information between multiple levels. However, these methods are susceptible to the influence of messy background information. The network may regard non-salient objects with similar salient appearances as target predictions, and the prediction results may be incomplete due to different appearance areas of salient objects. We design a network composed of multiple attention mechanisms to selectively integrate deep and shallow feature information, and more effectively deal with the transfer and fusion of features. In this paper, we propose a multi-attention embedded network (MAENet), which introduces attention mechanisms to give different feature information with different weights for handling the transfer and aggregation of features at different levels. The multi-attention feature aggregation (MAFA) module is proposed, which uses the channel attention mechanism to give different weights to the features to be fused, and then uses the spatial attention mechanism to selectively aggregate shallow edge information and deep abstract semantic features to avoid excessive redundant information which affects the saliency mapping, as well as suppressing non-salient areas with "salient" appearance. In addition, The multi-scale feature extraction (MFE) module and the self-attention (SA) module are also proposed for obtaining sufficiently rich and useful multi-scale context information and enhancing the function of the top layer. Finally, the attentional residual refinement (ARR) module is utilized to refine the saliency map after each feature fusion and further improve the input function. MAENet can accurately segment salient objects and provide clear local details. Experimental results on five benchmark datasets show that the proposed method achieves the favorable performance against 14 state-of-the-art methods on popular evaluation metrics.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1432-7643  
eISSN:  
1433-7479  
入藏号:  
WOS:000695782400001  
Indexed Date:  
2021-09-23  
Institution Address:  
China Jiliang Univ, Dept Informat Engn, Hangzhou 310018, Peoples R China  
Institution Address:  
Lishui Peoples Hosp, Lishui, Zhejiang, Peoples R China  
Institution:  
Institution:  
Lishui Peoples Hosp

### 492. PCTMF-Net: heart sound classification with parallel CNNs-transformer and

摘要: Heart disease is a common condition worldwide and has become one of the leading causes of death worldwide. The electrocardiogram (PCG) is a safe, painless, and non-invasive test that captures bioacoustic information reflecting the function of the heart by capturing the acoustic signal of the patient's heart. Nowadays, based on biosignal processing and artificial intelligence technologies, automated heart sound classification is playing an increasingly important role in clinical applications. In this paper, we propose a new parallel CNNs-transformer network with multi-scale feature context aggregation (PCTMF-Net). It combines the advantages of CNNs and transformer to achieve efficient heart sound classification. In PCTMF-Net, firstly, the heart tone signal features are extracted using the second-order spectral analysis, and a transformer-based MHTE-4 (multi-head transformer encoder with four attention heads) is designed to encode and aggregate the contextual information, and then, two CNNs feature extractors are designed in parallel with MHTE-4 to capture the hierarchical features. Finally, the feature vectors obtained from CNNs and MHTE-4 through feature fusion in PCTMF-Net will be fed into the fully connected layer for predicting the classification results of heart sounds. In addition, we perform validation based on two publicly available mutually exclusive heart sound datasets and conduct extensive experiments and comparisons of existing algorithms under different metrics. The experimental results show that our proposed method achieves 99.36% accuracy on the Yaseen dataset and 93% accuracy on the PhysioNet dataset. It surpasses current algorithms in terms of accuracy, recall and F1-score metrics. The aim of this study is to apply these new techniques and methods to improve the diagnostic accuracy and validity of heart disease for clinical use.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
10  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001032577400002  
Indexed Date:  
2023-08-04  
Institution Address:  
Macao Polytech Univ, Fac Appl Sci, Rua Luis Gonzaga Gomes, Taipa 999078, Macao, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Fac John Hopcroft Ctr, Shanghai 200240, Peoples R China  
Institution:  
Institution:

### 493. Single image deraining using multi-scales context information and

摘要: The existing deraining methods based on convolutional neural networks (CNNs) have made great success, but some remaining rain streaks can degrade images drastically. In this work, we proposed an end-to-end multi-scale context information and attention network, called MSCIANet. The proposed network consists of multi-scale feature extraction (MSFE) and multi-receptive fields feature extraction (MRFFE). Firstly, the MSFE can pick up features of rain streaks in different scales and propagate deep features of the two layers across stages by skip connections. Secondly, the MRFFE can refine details of the background by attention mechanism and the depthwise separable convolution of different receptive fields with different scales. Finally, the fusion of these outputs of two subnetworks can reconstruct the clean background image. Extensive experimental results have shown that the proposed network achieves a good effect on the deraining task on synthetic and real-world datasets. The demo can be available at https://github.com/CoderLi365/MSCIANet.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
1047-3203  
eISSN:  
1095-9076  
入藏号:  
WOS:000974535500001  
Indexed Date:  
2023-05-09  
Institution Address:  
Nanchang Hangkong Univ, Sch Informat Engn, Nanchang 330063, Jiangxi, Peoples R China  
Institution:

### 494. Adaptive convolutional neural network for aluminum surface defect

摘要: The aluminum surface defect detection is not trivial for the high computational cost and labor-intensive data annotation. Particularly, the characteristics of lots of very tiny objects, sample sparsity, and variations, limit the detection performance. In this paper, we demonstrate a sophisticated and efficient object detection model based on hierarchical attention and contextual information for aluminum surface defect detection. Specially, we first use a deep residual learning strategy to obtain the defect feature maps. Secondly, we add corresponding weight matrices to the defect feature maps by fusing attention mechanism and adaptive deformable convolution to achieve the fine feature. Thirdly, we construct a feature pyramid structure to achieve the fusion of multi-scale feature information. Finally, we use the obtained contextual feature information for class prediction and bounding box regression. The comprehensive experiments on the surface defect data set of aluminum and the surface defect data set of copper foil and aluminum foil respectively show that our method compared to state-of-the-art object detectors. Code is available at https://github.com/yunsheng-Wei/DFA-FRCNN.git.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
8  
ISSN:  
0927-0256  
eISSN:  
1879-0801  
入藏号:  
WOS:001009274100001  
Indexed Date:  
2023-06-29  
Institution Address:  
Hefei Univ, Sch Energy Mat & Chem Engn, Hefei 230601, Anhui, Peoples R China  
Institution Address:  
Hefei Univ, Inst Appl Optimizat, Sch Artificial Intelligence & Big Data, Hefei 230601, Anhui, Peoples R China  
Institution Address:  
Hefei Univ, Sch Artificial Intelligence & Big Data, Anhui Prov Engn Lab Big Data Technol Applicat Urba, Hefei 230601, Anhui, Peoples R China  
Institution:  
Institution:  
Institution:

### 495. DM-CNN: Dynamic Multi-scale Convolutional Neural Network with

摘要: Convolutional neural network (CNN) has promoted the development of diagnosis technology of medical images. However, the performance of CNN is limited by insufficient feature information and inaccurate attention weight. Previous works have improved the accuracy and speed of CNN but ignored the uncertainty of the prediction, that is to say, uncertainty of CNN has not received enough attention. Therefore, it is still a great challenge for extracting effective features and uncertainty quantification of medical deep learning models In order to solve the above problems, this paper proposes a novel convolutional neural network model named DM -CNN, which mainly contains the four proposed sub -modules : dynamic multi -scale feature fusion module (DMFF), hierarchical dynamic uncertainty quantifies attention (HDUQ-Attention) and multi -scale fusion pooling method (MF Pooling) and multi -objective loss (MO loss). DMFF select different convolution kernels according to the feature maps at different levels, extract different -scale feature information, and make the feature information of each layer have stronger representation ability for information fusion HDUQ-Attention includes a tuning block that adjust the attention weight according to the different information of each layer, and a Monte -Carlo (MC) dropout structure for quantifying uncertainty MF Pooling is a pooling method designed for multi -scale models, which can speed up the calculation and prevent overfitting while retaining the main important information Because the number of parameters in the backbone part of DM -CNN is different from other modules, MO loss is proposed, which has a fast optimization speed and good classification effect DM -CNN conducts experiments on publicly available datasets in four areas of medicine (Dermatology, Histopathology, Respirology, Ophthalmology), achieving state-of-the-art classification performance on all datasets. DM -CNN can not only maintain excellent performance, but also solve the problem of quantification of uncertainty, which is a very important task for the medical field. The code is available: https://github.com/QIANXIN22/DM-CNN.  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
18  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
19  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:001164673200001  
PubMed ID:  
38042102  
Indexed Date:  
2024-03-17  
Institution Address:  
Chongqing Univ Sci & Technol, Sch Intelligent Technol & Engn, Chongqing 401331, Peoples R China  
Institution Address:  
Chongqing Univ Sci & Technol, Sch Elect Engn, Chongqing 401331, Peoples R China  
Institution Address:  
Tarim Univ, Sch Informat Engn, Alar City 843300, Peoples R China  
Institution:  
Institution:  
Institution:

### 496. PDS-Net: A novel point and depth-wise separable convolution for

摘要: Numerous recent object detectors and classifiers have shown acceptable performance in recent years by using convolutional neural networks and other efficient architectures. However, most of them continue to encounter difficulties like overfitting, increased computational costs, and low efficiency and performance in real-time scenarios. This paper proposes a new lightweight model for detecting and classifying objects in images. This model presents a backbone for extracting in-depth features and a spatial feature pyramid network (SFPN) for accurately detecting and categorizing objects. The proposed backbone uses point-wise separable (PWS) and depth-wise separable convolutions, which are more efficient than standard convolution. The PWS convolution utilizes a residual shortcut link to reduce computation time. We also propose a SFPN that comprises concatenation, transformer encoder-decoder, and feature fusion modules, which enables the simultaneous processing of multi-scale features, the extraction of low-level characteristics, and the creation of a pyramid of features to increase the effectiveness of the proposed model. The proposed model outperforms all of the existing backbones for object detection and classification in three publicly accessible datasets: PASCAL VOC 2007, PASCAL VOC 2012, and MS-COCO. Our extensive experiments show that the proposed model outperforms state-of-the-art detectors, with mAP improvements of 2.4% and 2.5% on VOC 2007, 3.0% and 2.6% on VOC 2012, and 2.5% and 3.6% on MS-COCO in the small and large sizes of the images, respectively. In the MS-COCO dataset, our model achieves FPS of 39.4 and 33.1 in a single GPU for the small (320 x 320) and large (512 x 512) sizes of the images, respectively, which shows that our method can run in real-time.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
ISSN:  
2192-6611  
eISSN:  
2192-662X  
入藏号:  
WOS:000772700800001  
Indexed Date:  
2022-03-24  
Institution Address:  
Bahcesehir Univ, Dept Comp Engn, TR-34349 Istanbul, Turkey  
Institution:

### 497. Application of Tswin-F network based on multi-scale feature fusion in

摘要: Tomato leaf lesion identification can greatly help the detection and analysis of plant lesions. This study proposes Tswin-F network, a new network structure based on Transformer, to detect tomato leaf diseases. This Tswin-F network would obtain position information on images by implementing the bilateral local attention module and the self-supervised learning module. Specifically, the bilateral local attention mechanism focuses on the connection with certain continuous tokens, while the self-supervised learning module pays attention to the connection with random token positions. Then the information learned from the above two modules approaches will be combined to create the spatial connection between the final tokens. The combination of the above two modules can enhance the ability to communicate information between the windows of the input images and improve the accuracy of the models. In addition, a Feature Fuse Local Attention (FFLCA) structure is designed to solve the problem that attention distances would increase with the number of layers in the transformer network model. Furthermore, all the feature information is fused through the adaptive fusion strategy and is inputted into the classification network as the final global information of the model. Finally, an accuracy of 99.64% is obtained on 10 types of datasets, reaching the state-of-the-art level of CNN-based methods in terms of accuracy. The accuracy rate of identifying 13 types of tomato leaf lesions reaches 90.81% on average. Code is available at: htt ps://github.com/fightpotato.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:001274614600001  
Indexed Date:  
2024-07-28  
Institution Address:  
Zhejiang A&F Univ, Sch Math & Comp Sci, Hangzhou 311300, Peoples R China  
Institution Address:  
Zhejiang Prov Key Lab Forestry Intelligent Monitor, Hangzhou 311300, Peoples R China  
Institution Address:  
Zhejiang Univ, Coll Informat Sci & Elect Engn, Hangzhou 310027, Peoples R China  
Institution Address:  
State Key Lab CAD & CG, Hangzhou 310027, Peoples R China  
Institution:  
Institution:  
Zhejiang Prov Key Lab Forestry Intelligent Monitor  
Institution:  
Institution:  
State Key Lab CAD & CG

### 498. Fast instruments and tissues segmentation of micro-neurosurgical scene

摘要: Surgical scene segmentation provides critical information for guidance in micro-neurosurgery. Segmentation of instruments and critical tissues contributes further to robot assisted surgery and surgical evaluation. However, due to the lack of relevant scene segmentation dataset, scale variation and local similarity, micro-neurosurgical segmentation faces many challenges. To address these issues, a high correlative non-local network (HCNNet), is proposed to aggregate multi-scale feature by optimized non-local mechanism. HCNNet adopts two-branch design to generate features of different scale efficiently, while the two branches share common weights in shallow layers. Several short-term dense concatenate (STDC) modules are combined as the backbone to capture both semantic and spatial information. Besides, a high correlative non-local module (HCNM) is designed to guide the upsampling process of the high-level feature by modeling global context generated from the low-level feature. It filters out confused pixels of different classes in the non-local correlation map. Meanwhile, a large segmentation dataset named NeuroSeg is constructed, which contains 15 types of instruments and 3 types of tissues that appear in meningioma resection surgery. The proposed HCNNet achieves the state-of-the-art performance on NeuroSeg, it reaches an inference speed of 54.85 FPS with the highest accuracy of 59.62% mIoU, 74.7% Dice, 70.55% mAcc and 87.12% aAcc.  
在 BIOSIS Citation Index 中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0010-4825  
eISSN:  
1879-0534  
入藏号:  
WOS:000976578600001  
PubMed ID:  
36638619  
Indexed Date:  
2023-05-14  
Institution Address:  
Hebei Univ Technol, Sch Artificial Intelligence & Data Sci, Tianjin 300131, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Management & Control Complex Syst, Beijing 100190, Peoples R China  
Institution Address:  
Tongji Univ, Sch Elect & Informat Engn, Shanghai 201804, Peoples R China  
Institution Address:  
Capital Med Univ, Beijing Tiantan Hosp, Beijing 100190, Peoples R China  
Institution Address:  
Beijing Informat Sci & Technol Univ, Sch Automat, Beijing 100192, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automat, Beijing 100190, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 499. A medical image segmentation method for rectal tumors based on

摘要: BackgroundWith the continuous development of deep learning algorithms in the field of medical images, models for medical image processing based on convolutional neural networks have made great progress. Since medical images of rectal tumors are characterized by specific morphological features and complex edges that differ from natural images, achieving good segmentation results often requires a higher level of enrichment through the utilization of semantic features.PurposeThe efficiency of feature extraction and utilization has been improved to some extent through enhanced hardware arithmetic and deeper networks in most models. However, problems still exist with detail loss and difficulty in feature extraction, arising from the extraction of high-level semantic features in deep networks.MethodsIn this work, a novel medical image segmentation model has been proposed for Magnetic Resonance Imaging (MRI) image segmentation of rectal tumors. The model constructs a backbone architecture based on the idea of jump-connected feature fusion and solves the problems of detail feature loss and low segmentation accuracy using three novel modules: Multi-scale Feature Retention (MFR), Multi-branch Cross-channel Attention (MCA), and Coordinate Attention (CA).ResultsCompared with existing methods, our proposed model is able to segment the tumor region more effectively, achieving 97.4% and 94.9% in Dice and mIoU metrics, respectively, exhibiting excellent segmentation performance and computational speed.ConclusionsOur proposed model has improved the accuracy of both lesion region and tumor edge segmentation. In particular, the determination of the lesion region can help doctors identify the tumor location in clinical diagnosis, and the accurate segmentation of the tumor edge can assist doctors in judging the necessity and feasibility of surgery.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0094-2405  
eISSN:  
2473-4209  
入藏号:  
WOS:001196051900001  
PubMed ID:  
38569054  
Indexed Date:  
2024-04-12  
Institution Address:  
Taiyuan Univ Technol, Coll Informat & Comp, Jinzhong, Peoples R China  
Institution Address:  
Key Lab Big Data Fus Anal & Applicat Shanxi Prov, Taiyuan, Peoples R China  
Institution Address:  
Intelligent Percept Engn Technol Ctr Shanxi, Taiyuan, Peoples R China  
Institution Address:  
Shanxi Prov Engn Technol Res Ctr Spatial Informat, Taiyuan, Peoples R China  
Institution Address:  
Taiyuan Univ Technol, Coll Data Sci, Jinzhong 030600, Peoples R China  
Institution Address:  
Shanxi Med Univ, Shanxi Prov Canc Hosp, Dept Radiol, Taiyuan 030013, Peoples R China  
Institution:  
Institution:  
Key Lab Big Data Fus Anal & Applicat Shanxi Prov  
Institution:  
Intelligent Percept Engn Technol Ctr Shanxi  
Institution:  
Shanxi Prov Engn Technol Res Ctr Spatial Informat  
Institution:  
Institution:

### 500. A novel approach based on a modified mask R-CNN for the weight

摘要: Since determining the weight of pigs during large-scale breeding and production is challenging, using noncontact estimation methods is vital. This study proposed a novel pig weight prediction method based on a modified mask region-convolutional neural network (mask R-CNN). The modified approach used ResNeSt as the backbone feature extraction network to enhance the image feature extraction ability. The feature pyramid network (FPN) was added to the backbone feature extraction network for multi-scale feature fusion. The channel attention mechanism (CAM) and spatial attention mechanism (SAM) were introduced in the region proposal network (RPN) for the adaptive integration of local features and their global dependencies to capture global information, ultimately improving image segmentation accuracy. The modified network obtained a precision rate (P), recall rate (R), and mean average precision (MAP) of 90.33%, 89.85%, and 95.21%, respectively, effectively segmenting the pig regions in the images. Five image features, namely the back area, body length, body width, average depth, and eccentricity, were investigated. The pig depth images were used to build five regression algorithms (ordinary least squares (OLS), AdaBoost, CatBoost, XGBoost, and random forest (RF)) for weight value prediction. AdaBoost achieved the best prediction result with a coefficient of determination (R 2 ) of 0.987, a mean absolute error (MAE) of 2.96 kg, a mean square error (MSE) of 12.87 kg 2 , and a mean absolute percentage error (MAPE) of 8.45%. The results demonstrated that the machine learning models effectively predicted the weight values of the pigs, providing technical support for intelligent pig farm management. (c) 2024 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2589-7217  
入藏号:  
WOS:001225677300001  
Indexed Date:  
2024-05-24  
Institution Address:  
Zhejiang Acad Agr Sci, Inst Anim Husb & Vet Sci, State Key Lab Managing Biot & Chem Threats Qual &, Hangzhou 310021, Peoples R China  
Institution Address:  
China Jiliang Univ, Coll Informat Engn, Hangzhou 310018, Peoples R China  
Institution Address:  
Zhejiang Tongji Vocat Coll Sci & Technol, Coll Hydraul Engn, Hangzhou 311231, Peoples R China  
Institution:  
Institution:  
Institution:  
Zhejiang Tongji Vocat Coll Sci & Technol

### 501. A Multi-Organ Segmentation Network Based on Densely Connected RL-Unet

摘要: The convolutional neural network (CNN) has been widely applied in medical image segmentation due to its outstanding nonlinear expression ability. However, applications of CNN are often limited by the receptive field, preventing it from modeling global dependencies. The recently proposed transformer architecture, which uses a self-attention mechanism to model global context relationships, has achieved promising results. Swin-Unet is a Unet-like simple transformer semantic segmentation network that combines the dominant feature of both the transformer and Unet. Even so, Swin-Unet has some limitations, such as only learning single-scale contextual features, and it lacks inductive bias and effective multi-scale feature selection for processing local information. To solve these problems, the Residual Local induction bias-Unet (RL-Unet) algorithm is proposed in this paper. First, the algorithm introduces a local induction bias module into the RLSwin-Transformer module and changes the multi-layer perceptron (MLP) into a residual multi-layer perceptron (Res-MLP) module to model local and remote dependencies more effectively and reduce feature loss. Second, a new densely connected double up-sampling module is designed, which can further integrate multi-scale features and improve the segmentation accuracy of the target region. Third, a novel loss function is proposed that can significantly enhance the performance of multiple scales segmentation and the segmentation results for small targets. Finally, experiments were conducted using four datasets: Synapse, BraTS2021, ACDC, and BUSI. The results show that the performance of RL-Unet is better than that of Unet, Swin-Unet, R2U-Net, Attention-Unet, and other algorithms. Compared with them, RL-Unet produces significantly a lower Hausdorff Distance at 95% threshold (HD95) and comparable Dice Similarity Coefficient (DSC) results. Additionally, it exhibits higher accuracy in segmenting small targets.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001311036900001  
Indexed Date:  
2024-09-21  
Institution Address:  
Shanghai Inst Technol, Sch Elect & Elect Engn, Shanghai 201418, Peoples R China  
Institution:

### 502. A Lightweight Network for Crack Detection With Split Exchange

摘要: Crack detection mostly benefits from the rapid development of Convolutional Neural Networks (CNNs). However, the improvement of crack detection performance comes from the deeper and wider network structure, which requires heavier computation and storage overhead. This prevents crack detection methods from being deployed on practical platforms, especially mobile devices. To tackle this problem, we propose a novel Split Exchange Convolution (SEConv) modules, which splits the feature maps into high resolution and low resolution parts and then filters out the redundant information of each part. SEConv exchanges the feature information between the two modules to make the feature efficient reuse. Besides, we design a Multi-Scale Feature Exchange (MSFE) module to promote the cross stage features fusion. Benefiting from the SEConv and the MSFE modules, we build an extremely lightweight crack detection model with only 1.3 M parameters and 8 G FLOPs while achieving comparable performance. Extensive experimental results on the crack detection benchmark show that our method consistently outperforms other state-of-the-art methods in the evaluation metrics of F1-score and MIoU.  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
16  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
17  
ISSN:  
2379-8858  
eISSN:  
2379-8904  
入藏号:  
WOS:000981348100026  
Indexed Date:  
2023-06-12  
Institution Address:  
Chongqing Univ Posts & Telecommun, Sch Comp Sci & Technol, Chongqing 400065, Peoples R China  
Institution:

### 503. Remote Sensing Image Segmentation Based on Attention Guidance and

摘要: Objective Remote sensing images have a large detection range, long dynamic monitoring time, and a large amount of carrying information, making the obtained ground feature information more comprehensive and rich. By extracting ground object targets from remote sensing images, more detailed and accurate ground object information in the imaging area can be obtained, providing data support for high-altitude reconnaissance, precision guidance, and terrain matching. However, with the rapid increase in data volume, the current low level of intelligent and automated target extraction methods is difficult to embrace the demand. Traditional image extraction techniques contain edge detection, threshold segmentation, and region segmentation. These methods have good segmentation performance for remote sensing targets with significant contour boundaries but lack the ability of adaptive adjustment while facing complex and ever-changing remote sensing targets. Convolutional neural networks have stronger representation ability, scalability, and robustness than traditional methods by providing multi-level semantic information in images. Due to the uneven distribution, blurred edges, and variable scales of ground objects in remote sensing images, convolutional neural networks are prone to losing edge information and multi-scale feature information during feature extraction. In addition, cloud cover of remote sensing targets in complex scenes exacerbates the loss of target edge and multi-scale information, making it more difficult for convolutional neural networks to accurately segment remote sensing ground objects. In order to solve the above problems, we propose a segmentation method that uses deep residual networks as the backbone and combines attention guidance and multi-feature fusion to enhance the network's ability to segment remote sensing image ground object edges and multi-scale objects.  
Methods We propose a remote sensing image semantic segmentation network called AMSNet, which combines attention guidance and multi-feature fusion. In the Encoder Section, D\_ Resnet50 is applied as the backbone network to extract the main feature information from remote sensing images, which can enhance the acquisition of detailed information such as edge and small-scale targets in remote sensing images. The category guidance channel attention module is inserted into the backbone to enhance the network's segmentation ability for difficult-to-distinguish and irregularly shaped areas in remote sensing images. A feature reuse module is added to the backbone network to solve the loss of edge detail information and the disappearance of scattered small-scale targets during feature extraction. In the Decoder Section, the cross-regional feature fusion module is applied to fuse the multi-feature information, improving the acquisition of multi-scale target information. Multi-scale loss fusion module is also joined to further enhance the segmentation performance of the network for multi-scale targets.  
Results and Discussions From the analysis of experimental results on the remote sensing image dataset of the plateau region and the remote sensing image dataset of the plateau region under cloud interference, compared with other semantic segmentation networks, the proposed network has better segmentation performance (Table 6 and Table 7) regardless of cloud interference. In addition, the segmentation performance is less affected by cloud interference. Even under cloud interference, the segmentation accuracy of ground targets is only 1. 10 percentage points lower than that without cloud interference in mIoU, 0. 58 percentage points lower than that in mPa, and 0. 71 percentage points lower than that in mF1, which is lower than the influence of other semantic segmentation networks on segmentation effect under different cloud meteorological interference conditions. In addition, in order to verify the generalization performance of the AMSNet network segmentation effect, the International Society for Photogrammetry and Remote Sensing (ISPRS) dataset in the Vaihingen region of Germany is selected. In order to better fit the picture size, number of grouping convolutions of feature multiplexing modules in the AMSNet network is reduced to four groups. From the experimental results in Table 8, the network still performs better than other networks. This network is compared with PspNet and OCNet, with mIoU increased by 5. 09 percentage points and 5. 57 percentage points, Deeplabv3+ network with mIoU by 3. 47 percentage points, mPa by 3. 56 percentage points, and mF1 by 2. 78 percentage points. From the segmenting effect diagram of Fig. 8, this network has a lower error rate, fewer omission, and a more accurate segmenting boundary for building edges and small-scale cars than other networks.  
Conclusions We propose a network model based on encoding-decoding structure-AMSNet. In the encoding part, the D\_Resnet50 network is applied as the backbone to extract the main feature information of remote sensing images. We also use a category-guided channel attention module to reduce the interference of channel noise on segmented objects and improve the segmentation effect of targets in difficult-to-distinguish areas. We embed a feature reuse module to compensate for the problem of target edge loss and small-scale target loss during the feature extraction process. In the decoding part, the cross-regional feature fusion module is designed to integrate multi-layer features and combine the multiscale loss fusion module to calculate the feature loss at different scales to improve the segmentation effect of the network on multi-scale targets. This network conducts experiments on the remote sensing image dataset of the plateau region, remote sensing image dataset of the plateau region under cloud interference, and a public dataset. Compared with semantic segmentation networks such as BiseNetv2, PspNet, and Deeplabv3+, the proposed network achieves better results in the evaluation indicators of mIoU, mPa, and mF1. The visualization results show that the proposed network can effectively segment the ground object targets and scattered multi-scale targets in the interlaced and hard-to-distinguish areas in the remote sensing images, and it has good segmentation performance and good robustness in cloud interference.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
被引频次合计:  
10  
ISSN:  
0253-2239  
入藏号:  
WOS:001141747200034  
Indexed Date:  
2024-02-04  
Institution Address:  
Kunming Univ Sci & Technol, Fac Mech & Elect Engn, Kunming 650500, Yunnan, Peoples R China  
Institution Address:  
Rocket Force Engn Univ, Coll Missile Engn, Xian 710025, Shaanxi, Peoples R China  
Institution:  
Institution:

### 504. PS5-Net: a medical image segmentation network with multiscale resolution

摘要: Purpose: In recent years, the continuous advancement of convolutional neural networks (CNNs) has led to the widespread integration of deep neural networks as a mainstream approach in clinical diagnostic support. Particularly, the utilization of CNN-based medical image segmentation has delivered favorable outcomes for aiding clinical diagnosis. Within this realm, network architectures based on the U-shaped structure and incorporating skip connections, along with their diverse derivatives, have gained extensive utilization across various medical image segmentation tasks. Nonetheless, two primary challenges persist. First, certain organs or tissues present considerable complexity, substantial morphological variations, and size discrepancies, posing significant challenges for achieving highly accurate segmentation. Second, the predominant focus of current deep neural networks on single-resolution feature extraction limits the effective extraction of feature information from complex medical images, thereby contributing to information loss via continuous pooling operations and contextual information interaction constraints within the U-shaped structure. Approach: We proposed a five-layer pyramid segmentation network (PS5-Net), a multiscale segmentation network with diverse resolutions that is founded on the U-Net architecture. Initially, this network effectively leverages the distinct features of images at varying resolutions across different dimensions, departing from prior single-resolution feature extraction methods to adapt to intricate and variable segmentation scenarios. Subsequently, to comprehensively integrate feature information from diverse resolutions, a kernel selection module is proposed to assign weights to features across different dimensions, enhancing the fusion of feature information from various resolutions. Within the feature extraction network denoted as PS-UNet, we preserve the classical structure of the traditional U-Net while enhancing it through the incorporation of dilated convolutions. Results: PS5-Net attains a Dice score of 0.9613 for liver segmentation on the CHLISC dataset and 0.8587 on the ISIC2018 dataset for skin lesion segmentation. Comparative analysis with diverse medical image segmentation methodologies in recent years reveals that PS5-Net has achieved the highest scores and substantial advancements. Conclusions: PS5-Net effectively harnesses the rich semantic information available at different resolutions, facilitating a comprehensive and nuanced understanding of the input medical images. By capitalizing on global contextual connections, the network adeptly captures the intricate interplay of features and dependencies across the entire image, resulting in more accurate and robust segmentation outcomes. The experimental validation of PS5-Net underscores its superior performance in medical image segmentation tasks, offering promising prospects for enhancing diagnostic and analytical processes within clinical settings. These results highlight the potential of PS5-Net to significantly contribute to the advancement of medical imaging technologies and ultimately improve patient care through more precise and reliable image analysis.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2329-4302  
eISSN:  
2329-4310  
入藏号:  
WOS:001178213400014  
PubMed ID:  
38379775  
Indexed Date:  
2024-04-04  
Institution Address:  
Qingdao Univ Sci & Technol, Coll Informat Sci & Technol, Qingdao, Peoples R China  
Institution Address:  
Natl Univ Singapore, Coll Design & Engn, Singapore, Singapore  
Institution Address:  
Harbin Inst Technol, Sch Astronaut, Harbin, Peoples R China  
Institution:  
Institution:  
Institution:

### 505. Enhanced Wind Field Spatial Downscaling Method Using UNET Architecture

摘要: Before 2008, China lacked high-coverage regional surface observation data, making it difficult for the China Meteorological Administration Land Data Assimilation System (CLDAS) to directly backtrack high-resolution, high-quality land assimilation products. To address this issue, this paper proposes a deep learning model named UNET\_DCA, based on the UNET architecture, which incorporates a Dual Cross-Attention module (DCA) for multiscale feature fusion by introducing Channel Cross-Attention (CCA) and Spatial Cross-Attention (SCA) mechanisms. This model focuses on the near-surface 10-m wind field and achieves spatial downscaling from 6.25 km to 1 km. We conducted training and validation using data from 2020-2021, tested with data from 2019, and performed ablation experiments to validate the effectiveness of each module. We compared the results with traditional bilinear interpolation methods and the SNCA-CLDASSD model. The experimental results show that the UNET-based model outperforms SNCA-CLDASSD, indicating that the UNET-based model captures richer information in wind field downscaling compared to SNCA-CLDASSD, which relies on sequentially stacked CNN convolution modules. UNET\_CCA and UNET\_SCA, incorporating cross-attention mechanisms, outperform UNET without attention mechanisms. Furthermore, UNET\_DCA, incorporating both Channel Cross-Attention and Spatial Cross-Attention mechanisms, outperforms UNET\_CCA and UNET\_SCA, which only incorporate one attention mechanism. UNET\_DCA performs best on the RMSE, MAE, and COR metrics (0.40 m/s, 0.28 m/s, 0.93), while UNET\_DCA\_ars, incorporating more auxiliary information, performs best on the PSNR and SSIM metrics (29.006, 0.880). Evaluation across different methods indicates that the optimal model performs best in valleys, followed by mountains, and worst in plains; it performs worse during the day and better at night; and as wind speed levels increase, accuracy decreases. Overall, among various downscaling methods, UNET\_DCA and UNET\_DCA\_ars effectively reconstruct the spatial details of wind fields, providing a deeper exploration for the inversion of high-resolution historical meteorological grid data.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:001245240900001  
Indexed Date:  
2024-06-21  
Institution Address:  
Datong Meteorol Bur, Datong 037010, Peoples R China  
Institution Address:  
Natl Meteorol Informat Ctr, Beijing 100044, Peoples R China  
Institution Address:  
Shanxi Meteorol Informat Ctr, Taiyuan 030006, Peoples R China  
Institution Address:  
Yuncheng Meteorol Bur, Yuncheng 044000, Peoples R China  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Geog Sci, Nanjing 210044, Peoples R China  
Institution:  
Datong Meteorol Bur  
Institution:  
Institution:  
Shanxi Meteorol Informat Ctr  
Institution:  
Yuncheng Meteorol Bur  
Institution:

### 506. ITF-WPI: Image and text based cross-modal feature fusion model for

摘要: As one of the necessary cash crops in China and many other countries, wolfberry is parasitized by multiple pests, and its yield is highly susceptible to being affected. On the other hand, agricultural pest backgrounds are complex. When identifying them, single-modal models cannot utilize diverse data types across modalities, resulting in low identification accuracy and data utilization. Traditional unimodal identification models can no longer meet the needs of multimodal data development in agriculture. To overcome these challenges, the ITFWPI cross-modal feature fusion model is proposed, which consists of CoTN and ODLS for parallel processing of images and text, respectively. We incorporate the Transformer structure (CoT), which focuses on contextual feature extraction, into CoTN to make full use of the rich static and dynamic linear fusion contexts between adjacent keys and improve the 4-stage network of CoTN using Pyramid Squeezed Attention (PSA) to improve the extraction of multi-scale feature structure information and effectively promote the interaction of in-depth features with multi-scale spatial information. The ODLS network constructed by introducing 1D convolutional and bidirectional LSTM stacking has been shown to have more robust text feature acquisition than other advanced convolutional neural network-long short-term memory (CNN-LSTM) models from experimental results, with a 30% reduction in MACCs compared to the optimal model. The results showed that ITF-WPI performed well in accuracy, F1 score, model size, and MACCs with 97.98%, 93.19%, 52.20 MB, and 7.828 G compared to the classical state-of-the-art (SOTA) model, lightweight SOTA model and advanced Transformer neural network synthesis, respectively. The model has critical practical applications for promoting the development of crossmodal models in agriculture and research on wolfberry pest control and improving wolfberry yields. The code and dataset for this study will be posted on GitHub (https://github.com/wemindful/Cross-modal-pest-Identifying) as soon as the study is released, and new data will be updated in the future.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
26  
在 BIOSIS Citation Index 中的被引频次:  
5  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
被引频次合计:  
28  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:001058789800001  
Indexed Date:  
2023-09-14  
Institution Address:  
Agr Informat Inst CAAS, Natl Agr Sci Data Ctr, Beijing 100081, Peoples R China  
Institution Address:  
Chinese Acad Agr Sci, Natl Nanfan Res Inst Sanya, Sanya 572024, Peoples R China  
Institution Address:  
Satya Wacana Christian Univ, Dept Informat Technol, Salatiga, Indonesia  
Institution Address:  
Chinese Acad Agr Sci, Agr Informat Inst, 12 Zhongguancun South St, Beijing, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 507. BPAT-UNet: Boundary preserving assembled transformer UNet for ultrasound

摘要: Background and Objective: Accurate and efficient segmentation of thyroid nodules on ultrasound im-ages is critical for computer-aided nodule diagnosis and treatment. For ultrasound images, Convolutional neural networks (CNNs) and Transformers, which are widely used in natural images, cannot obtain sat-isfactory segmentation results, because they either cannot obtain precise boundaries or segment small objects. Methods: To address these issues, we propose a novel Boundary-preserving assembly Trans-former UNet (BPAT-UNet) for ultrasound thyroid nodule segmentation. In the proposed network, a Bound-ary point supervision module (BPSM), which adopts two novel self-attention pooling approaches, is de-signed to enhance boundary features and generate ideal boundary points through a novel method. Mean-while, an Adaptive multi-scale feature fusion module (AMFFM) is constructed to fuse features and chan-nel information at different scales. Finally, to fully integrate the characteristics of high-frequency local and low-frequency global, the Assembled transformer module (ATM) is placed at the bottleneck of the network. The correlation between deformable features and features-among computation is characterized by introducing them into the above two modules of AMFFM and ATM. As the design goal and even-tually demonstrated, BPSM and ATM promote the proposed BPAT-UNet to further constrain boundaries, whereas AMFFM assists to detect small objects. Results: Compared to other classical segmentation net-works, the proposed BPAT-UNet displays superior segmentation performance in visualization results and evaluation metrics. Significant improvement of segmentation accuracy was shown on the public thyroid dataset of TN3k with Dice similarity coefficient (DSC) of 81.64 % and 95th percentage of the asymmet-ric Hausdorff distance (HD95) of 14.06, whereas those on our private dataset were with DSC of 85.63 % and HD95 of 14.53, respectively. Conclusions: This paper presents a method for thyroid ultrasound im-age segmentation, which achieves high accuracy and meets the clinical requirements. Code is available at https://github.com/ccjcv/BPAT-UNet .& COPY; 2023 Elsevier B.V. All rights reserved.  
Web of Science 核心合集中的 "被引频次":  
27  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
7  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
28  
ISSN:  
0169-2607  
eISSN:  
1872-7565  
入藏号:  
WOS:001013279900001  
PubMed ID:  
37244233  
Indexed Date:  
2023-07-05  
Institution Address:  
Changzhou Univ, Sch Comp Sci & Artificial Intelligence, Changzhou 213164, Jiangsu, Peoples R China  
Institution Address:  
Nanjing Med Univ, Affiliated Changzhou Peoples Hosp NO 2, Changzhou 213003, Jiangsu, Peoples R China  
Institution Address:  
Southeast Univ, Key Lab Comp Network & Informat Integrat, Nanjing 211096, Jiangsu, Peoples R China  
Institution Address:  
Jiangsu Prov Engn Res Ctr Med Phys, Changzhou 213003, Jiangsu, Peoples R China  
Institution Address:  
Nanjing Med Univ, Ctr Med Phys, Changzhou 213003, Jiangsu, Peoples R China  
Institution Address:  
Changzhou Inst Technol, Changzhou 213032, Jiangsu, Peoples R China  
Institution Address:  
Southeast Univ, Lab Image Sci & Technol, Nanjing 210096, Jiangsu, Peoples R China  
Institution Address:  
Ctr Rech Informat Biomedicale Sino Francais, F-35000 Rennes, France  
Institution:  
Institution:  
Institution:  
Institution:  
Jiangsu Prov Engn Res Ctr Med Phys  
Institution:  
Institution:  
Institution:  
Institution:  
Ctr Rech Informat Biomedicale Sino Francais

### 508. Spikelets detection of table grape before thinning based on improved

摘要: Inflorescence thinning is the primary method for crop regulation to obtain high-quality table grapes in viticulture. It is essential to reduce labor dependency and associated costs by using mechanical thinning. To achieve precision mechanical thinning, visual detection of table grape inflorescence and spikelet is an important part of the process. In this article, an end-to-end method based on the improved YOLOV5s and Kmeans algorithm under the complex growing environment is proposed to detect spikelets needed to be removed from the table grape inflorescence. Firstly, the following improvements are made in the YOLOV5s: (1) The attention mechanism Pyramid Split Attention (PSA) establishes longer-distance channel dependencies. (2) The Bi-directional Feature Pyramid Network (BiFPN) enhances the multi-scale feature fusion. (3) CIoU loss makes more accurate regression of bounding box. Then, the test set is input into the improved YOLOV5s model to obtain the predicted bounding box of inflorescences and spikelets. The inflorescences are matched to the spikelets on them by the IoU function, and the Kmeans algorithm is used to cluster the center coordinates of the matched spikelet bounding boxes and determine the tail of the inflorescence according to the aggregation degree. 2/3 of the spikelet bounding boxes on the tail of inflorescence are taken as the spikelets removal. Finally, experiments are designed to verify the detection performance of the proposed method. What's more, compared with the original YOLOV5s, the improved model has 4.4 percentage points higher mAP value and 7ms slower detection speed than the original YOLOV5s, but still within the acceptable range. Compared with the Faster R-CNN, Cascade R-CNN, SSD, Retinanet, YOLOV3, and YOLOX-s, the improved YOLOV5s improves by 8, 7, 3.2, 9.6, 3.4, 1.9 percentage points in mAP, respectively. It indicates that the improved YOLOV5s detection accuracy and speed can reach a high level. In addition, the parameters of the proposed method are experimentally analyzed to determine the optimal parameters in this article. The results show that the algorithm has better accuracy when the number of spikelets threshold is 20, the IoU threshold is 0.15, the confidence threshold is 0.6, and its maximum accuracy is 78%. Therefore, the proposed algorithm has better detection accuracy and speed under the complex environment and provides theoretical support for the development of table grape thinning machinery.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
14  
在 BIOSIS Citation Index 中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
16  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:000882141100003  
Indexed Date:  
2022-11-24  
Institution Address:  
Shandong Agr Univ, Coll Mech & Elect Engn, Tai An 827101, Peoples R China  
Institution Address:  
Shandong Agr Equipment Intelligent Engn Lab, Tai An 827101, Peoples R China  
Institution Address:  
Shandong Prov Key Lab Hort Machinery & Equipment, Tai An 827101, Peoples R China  
Institution Address:  
Shandong Jiaotong Univ, Sch Construction Machinery, Jinan 250357, Peoples R China  
Institution:  
Institution:  
Shandong Agr Equipment Intelligent Engn Lab  
Institution:  
Shandong Prov Key Lab Hort Machinery & Equipment  
Institution:

### 509. Video object detection using fusion of SSD and spatiotemporal features

摘要: Objective Object detection is a fundamental task in computer vision applications,which provides support for subsequent object tracking,semantic segmentation,and behavior recognition.Recent years have witnessed substantial progress in still image object detection based on deep convolutional neural network (DCNN).The task of still image object detection is to determine the category and position of each object in an image.Video object detection aims to locate a moving object in sequential images and assign a specific category label to each object.The accuracy of video object detection suffers from degenerated object appearances in videos,such as motion blur,multiobject occlusion,and rare poses.The methods of still image object detection achieve excellent results,but directly applying them to video object detection is challenging.According to the temporal and spatial information in videos,most existing video object detection methods improve the accuracy of moving object detection by considering spatiotemporal consistency based on still image object detection.Method In this paper,we propose a video object detection method using fusion of single shot multibox detector (SSD) and spatiotemporal features.Under the framework of SSD,temporal and spatial information of the video are applied to video object detection through the optical flow network and the feature pyramid network.On the one hand,the network combining residual network (ResNet) 101 with four extra convolutional layers is used for feature extraction to produce the feature map in each frame of the video.An optical flow network estimates the optical flow fields between the current frame and multiple adjacent frames to enhance the feature of the current frame.The feature maps from adjacent frames are compensated to the current frame according to the optical flow fields.The multiple compensated feature maps as well the feature map of the current frame are aggregated according to adaptive weights.The adaptive weights indicate the importance of all compensated feature maps to the current frame.Here,the cosine similarity metric is utilized to measure the similarity between the compensated feature map and the feature map extracted from the current frame.If the compensated feature map is close to the feature map of the current frame,then the compensated feature map is assigned a larger weight;otherwise,it is assigned a smaller weight.Moreover,an embedding network that consists of three convolutional layers is applied on the compensated feature maps and the current feature map to produce the embedding feature maps,and the embedding feature maps are used to compute the adaptive weights.On the other hand,the feature pyramid network is used to extract multiscale feature maps that are used to detect the object of different sizes.The low-and high-level feature maps are used to detect smaller and larger objects,respectively.For the problem of small object detection in the original SSD network,the low-level feature map is combined with the high-level feature map to enhance the semantic information of the low-level feature map via upsampling operation and a 1 \* 1 convolutional layer.The upsampling operation is used to extend the high-level feature map to the same resolution as the low-level feature map,and the 1 \* 1 convolution layer is used to reduce the channel dimensions of the lowlevel feature map to be consistent with those of the high-level feature map.  
摘要:  
目的视频目标检测旨在序列图像中定位运动目标,并为各个目标分配指定的类别标签。视频目标检测存在目标模糊和多目标遮挡等问题,现有的大部分视频目标检测方法是在静态图像目标检测的基础上,通过考虑时空一致性来提高运动目标检测的准确率,但由于运动目标存在遮挡、模糊等现象,目前视频目标检测的鲁棒性不高。为此,本文提出了一种单阶段多框检测(single shot multibox detector,SSD)与时空特征融合的视频目标检测模型。方法在单阶段目标检测的SSD模型框架下,利用光流网络估计当前帧与近邻帧之间的光流场,结合多个近邻帧的特征对当前帧的特征进行运动补偿,并利用特征金字塔网络提取多尺度特征用于检测不同尺寸的目标,最后通过高低层特征融合增强低层特征的语义信息。结果实验结果表明,本文模型在ImageNet VID(Imagelvet for video object detetion)数据集上的mAP(mean average precision)为72.0%,相对于TCN(temporal convolutional networks)模型、TPN + LSTM(tubelet proposal network and long short term memory network)模型和SSD +孪生网络模型,分别提高了24.5%、3.6%和2.5%,在不同结构网络模型上的分离实验进一步验证了本文模型的有效性。结论本文模型利用视频特有的时间相关性和空间相关性,通过时空特征融合提高了视频目标检测的准确率,较好地解决了视频目标检测中目标漏检和误检的问题。  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1006-8961  
入藏号:  
CSCD:6930305  
Indexed Date:  
2021-06-12  
Institution Address:  
北京工业大学信息学部, 北京 100124, 中国  
Institution Address:  
Faculty of Information Technology,Beijing University of Technology, Beijing 100124, China  
Institution:  
北京工业大学信息学部  
Institution:  
Faculty of Information Technology,Beijing University of Technology

### 510. Combining Cascaded Network and Adversarial Network for Object Detection

摘要: Recognizing multi-scale objects and objects with occlusions is a key and difficult point of task in object detection.In order to detect objects with different sizes,the object detector usually uses the hierarchical structure of multi-scale feature map constructed by convolutional neural network (CNN).However,due to the small convolution layer of the bottom feature map,the top-down structure lacks the detailed information needed to capture the features of small object.The performance of these object detectors is limited.Therefore,based on the Faster R-CNN (regionconvolutional neural network) framework,this paper proposes Collaborative R-CNN.This paper designs a cascaded network structure that integrates multi-scale feature maps to generate deeply fused feature information and thereby improving the ability to detect small objects.Moreover,the quantization in the RoIPooling process greatly limits the recognition ability of small objects.In order to further improve the robustness of the method,a multi-scale RoIAlign is designed to eliminate such quantization,and the ability of network to detect objects with different scales is improved by multi-scale pooling.Finally,this paper combines an adversarial network with the proposed network to generate training samples with occlusions,significantly improving the classification ability of the model,and robustness to detect occlusions.Experimental results for the PASCAL VOC 2012 and PASCAL VOC 2007 datasets demonstrate the superiority of proposed approach relative to several state-of-the-art approaches.  
摘要:  
识别多尺度目标和遮挡目标是目标检测中的重点和难点。为了检测不同大小的目标,目标检测器通常利用卷积神经网络(CNN)的多尺度特征图层次结构,然而这种自顶向下的结构由于底层特征图的卷积层较小,缺乏获取小目标特征所需的细节信息,这些目标检测器的性能受到了限制。为此,结合Faster R-CNN框架提出Collaborative R-CNN,设计了一种级联网络结构,可以融合多尺度特征图,以生成深度融合的特征信息来增强小目标所需的细节特征,从而提高检测小目标的能力。此外,由于使用RoIPooling过程中的量化会对小目标检测造成极大的限制,为进一步提高方法的鲁棒性,设计了多尺度RoIAlign来消除这种量化,并通过多尺度的池化来提高网络检测不同尺度目标的能力。最后,将对抗网络与所提出的级联网络相结合,生成包含遮挡目标的训练样本,可显著提高模型的分类能力和识别遮挡目标的鲁棒性。在PASCAL VOC 2012和PASCAL VOC 2007数据集上的实验结果表明,提出的方法优于许多先进的方法。  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1673-9418  
入藏号:  
CSCD:7143424  
Indexed Date:  
2022-04-08  
Institution Address:  
广西师范大学, 广西多源信息挖掘与安全重点实验室, 桂林, 广西 541004, 中国  
Institution Address:  
西北师范大学计算机科学与工程学院, 兰州, 甘肃 730070, 中国  
Institution Address:  
Guangxi Normal University, Guangxi Key Lab of Multi-source Information Mining and Security, Guilin, Guangxi 541004, China  
Institution Address:  
College of Computer Science and Engineering,Northwest Normal University, Lanzhou, Gansu 730070, China  
Institution:  
广西师范大学  
Institution:  
西北师范大学计算机科学与工程学院  
Institution:  
Guangxi Normal University  
Institution:  
College of Computer Science and Engineering,Northwest Normal University

### 511. BDANet: Multiscale Convolutional Neural Network With Cross-Directional

摘要: Fast and effective responses are required when a natural disaster (e.g., earthquake and hurricane) strikes. Building damage assessment from satellite imagery is critical before relief effort is deployed. With a pair of predisaster and postdisaster satellite images, building damage assessment aims at predicting the extent of damage to buildings. With the powerful ability of feature representation, deep neural networks have been successfully applied to building damage assessment. Most existing works simply concatenate predisaster and postdisaster images as input of a deep neural network without considering their correlations. In this article, we propose a novel two-stage convolutional neural network for building damage assessment, called BDANet. In the first stage, a U-Net is used to extract the locations of buildings. Then, the network weights from the first stage are shared in the second stage for building damage assessment. In the second stage, a two-branch multiscale U-Net is employed as the backbone, where predisaster and postdisaster images are fed into the network separately. A cross-directional attention module is proposed to explore the correlations between predisaster and postdisaster images. Moreover, CutMix data augmentation is exploited to tackle the challenge of difficult classes. The proposed method achieves state-of-the-art performance on a large-scale dataset--xBD. The code is available at https://github.com/ShaneShen/BDANet-Building-Damage-Assessment.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
4  
Web of Science 核心合集中的 "被引频次":  
62  
在 BIOSIS Citation Index 中的被引频次:  
3  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
65  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:000732752500001  
Indexed Date:  
2021-12-28  
Institution Address:  
Minist Nat Resources, Inst Oceanog 2, State Key Lab Satellite Ocean Environm Dynam, Hangzhou 310012, Peoples R China  
Institution Address:  
Nanjing Univ Sci & Technol, Sch Comp Sci & Engn, Nanjing 210094, Peoples R China  
Institution Address:  
Southern Marine Sci & Engn Guangdong Lab, Guangzhou 511458, Peoples R China  
Institution Address:  
Univ North Carolina Charlotte, Dept Elect & Comp Engn, Charlotte, NC 28223 USA  
Institution Address:  
Mississippi State Univ, Dept Elect & Comp Engn, Starkville, MS 39762 USA  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 512. Ghost-guided UNet++ for high-resolution remote sensing image change

摘要: Objective With the rapid development of remote sensing observation technology, the resolutions of remote sensing images (RSIs) are increasing. Thus, how to extract discriminative features effectively from high-resolution RSIs for ground-object change detection has become a challenging problem. The existing RSI change detection methods can be divided into two categories: methods based on conventional image processing approaches and methods based on deep learning (DL) theory. The former extracts low-level or mid-level features from RSIs for change detection, making it easy to implement and have high detection efficiency. However, the increasing resolution of RSIs result in the images having rich ground objects and complex background clutter; thus, the low- or mid-level features can hardly meet the demand of precise change detection. In recent years, DL has been introduced into the field of high-resolution RSI change detection because of its powerful feature extraction capability. Various methods based on convolutional neural networks (CNNs) have been proposed for RSI change detection. Compared with conventional image processing methods, CNNs can extract high-level semantic information for high-resolution RSIs, which is beneficial to precise detection. Although CNNs have greatly raised the accuracy of change detection, they always involve numerous parameters and have high computational complexity. To raise the efficiency of change detection, many scholars have proposed to perform parameter pruning on pretrained models or design simple network structures. However, these strategies lead to the loss of some crucial image information, including semantics and location information, thus reducing the detection accuracy. Therefore, this study proposes a novel Ghost- UNet++ (GUNet++) network for precise RSI change detection to address the problems. Method First, a high-resolution network called HRNet, which has a multibranch architecture, is designed to replace the traditional UNet++ backbone and thus extract additional discriminative deep features from bitemporal RSIs. In contrast to series structures, HRNet owns a special parallel architecture, which can extract additional discriminative features through multiscale feature fusion. In addition, we choose a lightweight structure (i. e., HRNet-W16) on the basis of a thorough analysis of various existing HRNet structures to ensure that the whole network possesses low complexity. Second, when applying the UNet++ decoding structure for difference discrimination, the Ghost module is introduced to replace the conventional convolutional module and thus reduce the network parameters; meanwhile, a dense skip connection is designed to enhance the information transmission further and reduce the loss of location information. The core idea of the Ghost module is to adopt simple linear operations instead of the traditional convolutional operations to generate Ghost maps for original features, which may save substantial computational cost. Third, an ensemble attention module is constructed to aggregate and refine the multilevel semantic features of the network, thereby suppressing the loss of semantic and location information and further enhancing the feature representation ability for final accurate change detection. Features generated at various levels usually contain different meanings: shallow ones always contain detailed spatial information, while deep ones reflect rich semantic content. On this basis, we propose an adaptive channel selection mechanism to integrate these different features effectively. Finally, we propose to combine two different loss functions, i. e., the sigmoid loss function and the dice loss function, for the whole model training to enhance the detection performance further. Compared with the methods that merely use one loss function, this scheme can improve detection performance.  
摘要:  
目的随着遥感观测技术的飞速发展,遥感影像的分辨率越来越高,如何从高分遥感影像中有效提取具有鉴别性的特征进行地物变化检测成为一个具有挑战性的问题。卷积神经网络广泛应用于计算机视觉领域,但面向遥感影像变化检测时仍存在图像语义或位置信息的丢失及网络参数量过大等缺陷,导致检测性能受限。为此,提出一种新型GUNet++(Ghost-UNet++)网络,用于遥感影像的精准变化检测。方法首先,为了提取双时相遥感影像更具判别性的深度特征,设计具有多分支架构的高分辨率网络HRNet替换传统UNet++的主干网;其次,采用UNet++解码结构进行差异判别时,引入鬼影(Ghost)模块代替传统卷积模块以降低网络参量,并设计密集跳跃连接进一步加强信息传输,以减少深层位置信息的丢失;最后,设计一个集成注意力模块,将网络的多个语义层次特征进行聚合和细化,抑制语义和位置信息的丢失,进一步增强特征表征能力用于最终的精准变化检测。结果在LEVIR-CD (LEVIR change detection data set)和Google Data Set两个公开数据集上进行实验,结果表明本文算法变化检测精度高达99.62%和99.16%,且网络参数量仅为1.93 M,与现有主流变化检测方法相比优势明显。结论提出方法综合考虑了遥感图像中语义和位置信息对变化检测性能的影响,具有良好的特征抽取和表征能力,因此变化检测的精度和效率比现有同类方法更高。  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1006-8961  
入藏号:  
CSCD:7725016  
Indexed Date:  
2024-08-02  
Institution Address:  
河海大学计算机与信息学院, 南京, 江苏 211100, 中国  
Institution Address:  
School of Computer and Information,Hohai University, Nanjing, Jiangsu 211100, China  
Institution:  
河海大学计算机与信息学院  
Institution:  
School of Computer and Information,Hohai University

### 513. SANet-SI: A new Self-Attention-Network for Script Identification in

摘要: Developing an automatic method for identifying scripts in natural scene text images is of great impor-tance for improving performance of multilingual OCR. This paper presents a new Self-Attention Network (SANet-SI) for script identification in natural scene text images. The rationale behind proposing SANet-SI is that each script exhibits its own pattern because of different characteristics of scripts. To extract such observations, we explore self-attention-based CNN with a multi-scale feature extraction approach. The proposed multi-scale feature extraction involves local, global features extraction and fusion of both the features. Furthermore, to extract dominant features from the pool of features that contribute more for script identification, we explore Style-based Recalibration Module (SRM) in a new way. In addition, to improve the performance of the identification and reduce the model size, the proposed model uses the Global Average Pooling (GAP) layer, instead of Fully Connected(FC) layers in this work. The proposed model is evaluated on standard datasets, namely, RRC-MLT2017, SIW-13, and CVSI2015 to show effective-ness over state-of-the-art methods in terms of confusion matrix and classification rate. In addition, we also conducted experiments for Cross Dataset Validation to show that the proposed model is independent of the number of scripts and different datasets.(c) 2023 Elsevier B.V. All rights reserved.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
0167-8655  
eISSN:  
1872-7344  
入藏号:  
WOS:001001643100001  
Indexed Date:  
2023-06-16  
Institution Address:  
East China Normal Univ, Sch Elect & Comp Engn, Shanghai 200241, Peoples R China  
Institution Address:  
East China Normal Univ, Chongqing Inst, Chongqing 401120, Peoples R China  
Institution Address:  
Univ Malaya UM, Fac Comp Sci & Informat Technol, Kuala Lumpur 50603, Malaysia  
Institution Address:  
Indian Statict Inst, CVPR Unit, Kolkata 700108, India  
Institution:  
Institution:  
Institution:  
Institution:  
Indian Statict Inst

### 514. Prior-Experience-Based Vision-Language Model for Remote Sensing

摘要: Remote sensing (RS) image-text retrieval (RSITR) aims to retrieve relevant texts (RS images) based on the content of a given RS image (text). Existing methods are used to employing the convolutional neural network (CNN) and recurrent neural network (RNN) as encoders to learn visual and textual features for retrieval. Although feasible, the global information hidden in different data does not receive the attention it deserves. To mitigate this problem, transformers have been introduced. Nevertheless, the complexity of RS images present challenges in directly introducing Transformer-based architectures to multimodal learning in RS scenes, particularly in visual feature extraction and cross-modal interaction. In addition, the textual captions are always simpler than the complex RS images, leading to a semantic description appearing in different images. This typical false-negative (FN) sample problem increases the difficulty of RSITR tasks. To address the above limitations, we propose a new RSITR model named prior-experience-based RS vision-language (PERSVL). First, the specific visual and text encoders are used to extract features from RS images and texts. Also, a high-level feature complement (HFC) module is developed based on the self-attention mechanism (SAM) for the visual encoder to explore the complex contents from RS images fully. Second, a dual-branch multimodal fusion encoder (DBMFE) is designed to complete the cross-modal learning. It comprises a dual-branch multimodal interaction (DBMI) module and a branch fusion module. DBMI is designed to fully explore the relationships between different modalities, enriching visual and textual features. The branch fusion module integrates the cross-modal features and utilizes a classification head to generate matching scores for retrieval. Finally, a learning from prior experiences (LPEs) module is designed to reduce the influence of FN samples by analyzing the historical data produced in the model training process. Experiments are conducted on three popular datasets, and the positive results show that our PERSVL model achieves superior performance compared with previous methods. By integrating the advantages of natural language and RS images, our PERSVL can be applied in various applications, such as environmental monitoring, disaster evaluation, and urban planning. Our source codes are available at: https://github.com/TangXu-Group/Cross-modal-remote-sensing-image-and-text-retrieval-models/tree/main/PERSVL.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001329087800053  
Indexed Date:  
2024-10-19  
Institution Address:  
Xidian Univ, Minist Educ, Sch Artificial Intelligence, Key Lab Intelligent Percept & Image Understanding, Xian 710071, Peoples R China  
Institution Address:  
Nanjing Univ Sci & Technol, Minist Educ, Sch Comp Sci & Engn, Key Lab Intelligent Percept & Systemsfor High Dime, Nanjing 210094, Peoples R China  
Institution:  
Institution:

### 515. Enhanced Dual-Channel Model-Based with Improved Unet plus plus Network

摘要: Landslide disasters pose significant threats to human life and property; therefore, accurate and effective detection and area extraction methods are crucial in environmental monitoring and disaster management. In our study, we address the critical tasks of landslide detection and area extraction in remote sensing images using advanced deep learning techniques. For landslide detection, we propose an enhanced dual-channel model that leverages EfficientNetB7 for feature extraction and incorporates spatial attention mechanisms (SAMs) to enhance important features. Additionally, we utilize a deep separable convolutional neural network with a Transformers module for feature extraction from digital elevation data (DEM). The extracted features are then fused using a variational autoencoder (VAE) to mine potential features and produce final classification results. Experimental results demonstrate impressive accuracy rates of 98.92% on the Bijie City landslide dataset and 94.70% on the Landslide4Sense dataset. For landslide area extraction, we enhance the traditional Unet++ architecture by incorporating Dilated Convolution to expand the receptive field and enable multi-scale feature extraction. We further integrate the Transformer and Convolutional Block Attention Module to enhance feature focus and introduce multi-task learning, including segmentation and edge detection tasks, to efficiently extract and refine landslide areas. Additionally, conditional random fields (CRFs) are applied for post-processing to refine segmentation boundaries. Comparative analysis demonstrates the superior performance of our proposed model over traditional segmentation models such as Unet, Fully Convolutional Network (FCN), and Segnet, as evidenced by improved metrics: IoU of 0.8631, Dice coefficient of 0.9265, overall accuracy (OA) of 91.53%, and Cohen's kappa coefficient of 0.9185 on the Bijie City landslide dataset; and IoU of 0.8217, Dice coefficient of 0.9021, overall accuracy (OA) of 96.68%, and Cohen's kappa coefficient of 0.8835 on the Landslide4Sense dataset. These findings highlight the effectiveness and robustness of our proposed methodologies in addressing critical challenges in landslide detection and area extraction tasks, with significant implications for enhancing disaster management and risk assessment efforts in remote sensing applications.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:001305364800001  
Indexed Date:  
2024-09-09  
Institution Address:  
Beijing Normal Univ, Fac Arts & Sci, Zhuhai 519087, Peoples R China  
Institution Address:  
Beijing Normal Univ, Fac Geog Sci, State Key Lab Remote Sensing Sci, Beijing 100875, Peoples R China  
Institution:  
Institution:

### 516. A Multi- scale Hierarchical Residual Network- based Method for Tiny

摘要: Optical remote sensing image object detection aims to precisely locate and categorize targets such as aircraft, vehicles, and ships. Challenges arise due to the vast distances in remote sensing, leading to numerous tiny objects that are hard to characterize. Additionally, complex backgrounds and environmental factors like lighting and weather conditions reduce signal-to-noise ratios, increasing detection difficulties. Although Convolutional Neural Networks (CNNs), especially those from the YOLO family, are employed for their efficient feature extraction capabilities, they perform poorly in detecting these tiny objects. The key to realize the detection of tiny objects in optical remote sensing images is to obtain sufficiently rich multi-scale feature information and clear tiny object features. Aiming at the above problems, this paper proposes a multi-scale hierarchical residual network based optical remote sensing image tiny object detection algorithm MHRM-YOLO on the basis of YOLOv5, and designs a simple and efficient Multi-scale Hierarchical Residual tiny object feature extraction Module (MHRM). This module expands on Cross Stage Partial (CSP) module by doing more layered design and using different convolutional combinations to extract features from different layered, which allows the network to obtain richer gradient information flow and output richer feature map combinations. In addition, MHRM can be easily embedded into the existing mainstream YOLO detection algorithm backbone network, which can obtain richer sensory fields at a finer granularity level and can effectively capture the contextual information of tiny objects and retain their spatial feature information. The network structure of the MHRM-YOLO algorithm is mainly divided into three parts, namely the backbone, the neck, and the head for prediction. The backbone consists of MHRM and basic convolution module, which performs finegrained feature extraction to obtain more multi-scale information and larger sensory field; the neck part uses the conventional CSP plus Path Aggregation Network (PAN) feature pyramid structure to perform multi-scale feature fusion; and the prediction part uses the optimized localization loss function to perform computation. Since tiny object detection is sensitive to positional offsets during regression, the localization loss penalty term is further improved to enhance the algorithm's ability to perceive positional offsets. The shape penalty term of the baseline CIoU localization loss has lost its effect, in this regard, the optimized loss function retains the Euclidean distance penalty term of the centroid and adjusts it to a scalable exponential function, and improves the shape penalty term to a bounding box distance penalty term, which weakens the detection algorithm's sensitivity to positional offsets, and further improves the performance of the detection algorithm. In order to validate the effectiveness of the proposed detection algorithm, MHRM-YOLO conducts systematic experiments on the challenging optical remote sensing image tiny object detection dataset AITODv2 and the tiny pedestrian dataset TinyPerson. Systematic ablation experiments are conducted for the effects between different module combinations, the effects of the loss function, the performance difference between different backbone network modules and the portability of the algorithm, and the experimental results show that both the MHRM module and the localization loss function can improve the performance of the detection algorithm.  
Compared with the benchmark YOLOv5 algorithm, the average detection accuracy of MHRM-YOLO on the two datasets is improved by 5.5 degrees o and 1.8 degrees o respectively, which effectively reduces the false detection rate and the leakage rate of the detection of tiny objects in optical remote sensing images. Of course, due to the use of larger-scale feature layers for detection, the MHRMYOLO detection algorithm has an increased computational volume and a slight decrease in inference speed compared with the benchmark algorithm. The algorithm still has the problem of missed detection for relatively irregularly shaped target algorithms. In addition, the experimental results show that although the detection accuracy of the MHRM-YOLO algorithm has an advantage over the mainstream detection algorithms, the detection results are generally low, much lower than the accuracy of conventional target detection, and the algorithm still has room for further optimization.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1004-4213  
入藏号:  
WOS:001347674700001  
Indexed Date:  
2024-11-08  
Institution Address:  
Guangdong Univ Technol, Inst Adv Photon Technol, Sch Informat Engn, Key Lab Photon Technol Integrated Sensing & Commun, Guangzhou 510006, Peoples R China  
Institution Address:  
Guangdong Univ Technol, Inst Adv Photon Technol, Sch Informat Engn, Guangdong Prov Key Lab Informat Photon Technol, Guangzhou 510006, Peoples R China  
Institution Address:  
Southern Marine Sci & Engn Guangdong Lab Zhuhai, Zhuhai 519082, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Minist Ind & Informat Technol, Key Lab Light Field Manipulat & Informat Acquisit, Xian 710129, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Shaanxi Key Lab Opt Informat Technol, Sch Phys Sci & Technol, Xian 710129, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 517. SC-Net: Symmetrical conical network for colorectal pathology image

摘要: Background and Objective: Image segmentation of histopathology of colorectal cancer is a core task of computer aided medical image diagnosis system. Existing convolutional neural networks generally extract multi-scale information in linear flow structures by inserting multi-branch modules, which is difficult to extract heterogeneous semantic information under multi-level and different receptive field and tough to establish context dependency among different receptive field features.  
Methods: To address these issues, we propose a symmetric spiral progressive feature fusion encoder-decoder network called the Symmetric Conical Network (SC-Net). First, we design a Multi-scale Feature Extraction Block (MFEB) matching with the Symmetric Conical Network to obtain multi-branch heterogeneous semantic information under different receptive fields, so as to enrich the diversity of extracted feature information. The encoder is composed of MFEB through spiral and multi-branch arrangement to enhance context dependence between different information flow. Secondly, the information loss of contour, color and others in high-level semantic information through causally stacking MFEB, the Feature Mapping Layer (FML) is designed to map low-level features to high-level semantic features along the down-sampling branch and solve the problem of insufficient global feature extraction in deep levels.  
Results: The SC-Net was evaluated on our self-constructed colorectal cancer dataset, a publicly available breast cancer dataset and a polyp dataset. The results revealed that the mDice of segmentation reached 0.8611, 0.7259 and 0.7144. We compare our model with the state-of-art semantic segmentation UNet++, PSPNet, Attention U-Net, R2U-Net and other advanced segmentation networks. The experimental results demonstrate that we achieve the most advanced performance.  
Conclusions: The results indicate that the proposed SC-Net excels in segmenting H & E stained pathology images, effectively preserving morphological features and spatial information even in scenarios with weak texture, poor contrast, and variations in appearance.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0169-2607  
eISSN:  
1872-7565  
入藏号:  
WOS:001217734200001  
PubMed ID:  
38520785  
Indexed Date:  
2024-05-18  
Institution Address:  
Kunming Univ Sci & Technol, Fac Mech & Elect Engn, Kunming 650500, Peoples R China  
Institution Address:  
Kunming Med Univ, Yunnan Canc Hosp, Yunnan Canc Ctr, Dept Radiol,Affiliated Hosp 3, Kunming 650118, Peoples R China  
Institution:  
Institution:

### 518. A novel sEMG-based dynamic hand gesture recognition approach via

摘要: With the emergence of more andmore lightweight, convenient and cheap surface electromyography signal (sEMG) snsors, gesture recognition based on sEMG sensors has attracted much attention of researchers. In this study, combined with the sEMGsensor, a novel dynamic hand gesture recognition approach is proposed for effective and accurate dynamic gesture prediction. Here, a portable sEMG sensor (Myo wristband) is adopted to acquire the multi-channel sEMG signals of dynamic hand gestures and the continuous wavelet transformation (CWT) is proposed for data preprocessing to acquire the time-frequency maps. Due to the success of powerful contextual feature representation capability of deep convolutional neural networks (DCNNs), a deep residual attention network is proposed for accurate prediction of time-frequency maps. To effectively extract the key spatial and channel features from multi-channel sEMG signals, a residual attention network is proposed to act as the backbone network for effective feature representation. Besides, In the proposed recognition network, a multi-scale feature enhancement (MFE) module and an attention fusion block (AFB) are proposed, which respectively improve the multi-scale expression ability of the network and effectively realize multi-scale feature enhancement, respectively. Experimental results show that the proposed recognition network could achieve a superior detection ability compared with other state-of-the-art recognition models. The source code and dataset are available at https:// github.com/lyangucas92/ Ges\_Net.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001012579700005  
Indexed Date:  
2023-07-11  
Institution Address:  
Zhengzhou Univ, Sch Elect & Informat Engn, Zhengzhou 450001, Henan, Peoples R China  
Institution Address:  
Root Percept & Control Engn Lab, Zhengzhou 450001, Henan, Peoples R China  
Institution Address:  
Edinburgh Napier Univ, Built Environm, Edinburgh EH10 5DT, Scotland  
Institution:  
Institution:  
Root Percept & Control Engn Lab  
Institution:

### 519. U-Net segmentation model of brain tumor MR image based on attention

摘要: Objective Magnetic resonance (MR) image segmentation of brain tumors is crucial for patient evaluation and treatment.In recent years,as the feature extraction capability of convolutional neural networks improved,deep learning technology has been applied to medical image segmentation and achieved better results than those of traditional segmentation methods.The accuracy of image semantic segmentation depends on the effect of semantic feature extraction and processing.Traditional U-Net integrates high-level and low-level features inefficiently,leading to a loss of effective image information.In addition,U-Net does not make full use of context information.This study proposes a segmentation method for brain tumor MR images based on the attention mechanism and multi-view fusion U-Net algorithm.Method The model is modified as follows.First,in order to improve the network structure,residual structure is added to the U-Net to enhance the compensation of the network low-level information to the high-level information.On the one hand,the application of the residual idea,removes the same main part of the output of each convolution layer,highlights the small changes,makes each layer of the network more sensitive to the changes of the output,and makes the training easier,on the other hand,it alleviates the problem of gradient disappearance.Second,attention mechanism is added to the cascade structure,and the weight of tumor region is increased adaptively to enhance the utilization of effective information and improve the segmentation accuracy.Third,multi-scale feature fusion module is used to replace the traditional convolution layer between the down sampling structure and up sampling structure,and hole convolution with different sampling rates is used to extract and fuse the multiscale feature information of the image in parallel.The above is the adjustment of the network structure,and the influence of loss function and 3D structure of Brain tumor MR data set on model training is also considered in the experiment.There is a class imbalance problem in brain tumor segmentation,tumor area is smaller than normal brain tissue,network training is easily guided by a large number of irrelevant pixels,and linear combination of generalized dice loss and cross entropy loss is used to solve class imbalance problem and accelerate convergence.At the same time,Considering the influence of different view slices on the segmentation performance of the model,the three view slices of 3D MR image are trained respectively,and the multi view training model is fused in the segmentation prediction to improve the segmentation performance of the model.Result The proposed model is validated using the brain tumor MR image data set provided by Multi- modal Brain Tumor Segmentation Challenge 2018 (BraTS18),which includes four kinds of MR images of 210 high-grade gliomas(HGG) patients and 75 low-grade gliomas (LGG) patients and their real segmentation labels.The Dice similarity coefficient and Hausdorff distance95 are selected as technical indicators to further evaluate the accuracy of the brain tumor segmentation results.The effectiveness of the proposed module is proved by ablation experiments.After adding all modules,the performance of the model is optimal,the Dice scores of the entire tumor area,the core region,and the enhanced region reach 0.883,0.812,and 0.774.Compared with the traditional U-Net,it is improved by 3.9%,5.1% and 3.3% respectively,especially in the core region.After the fusion of three perspective slice training,comprehensive experiments show that the Dice scores of the entire tumor area,the core region,and the enhanced region reach 0.907,0.838,and 0.819 respectively,this algorithm exhibits better performance in terms of Dice score and Hausdorff 95 distance than others.  
摘要:  
目的脑肿瘤核磁共振(magnetic resonance,MR)图像分割对评估病情和治疗患者具有重要意义。虽然深度卷积网络在医学图像分割中取得了良好表现,但由于脑胶质瘤的恶性程度与外观表现有巨大差异,脑肿瘤MR图像分割仍是一项巨大挑战。图像语义分割的精度取决于图像特征的提取和处理效果。传统的U-Net网络以一种低效的拼接方式集成高层次特征和低层次特征,从而导致图像有效信息丢失,此外还存在未能充分利用上下文信息和空间信息的问题。对此,本文提出一种基于注意力机制和多视角融合U-Net算法,实现脑肿瘤MR图像的分割。方法在U-Net的解码和编码模块之间用多尺度特征融合模块代替传统的卷积层,进行多尺度特征映射的提取与融合;在解码模块的级联结构中添加注意力机制,增加有效信息的权重,避免信息冗余;通过融合多个视角训练的模型引入3维图像的空间信息。结果提出的模型在BraTS18(Multimodal Brain Tumor Segmentation Challenge 2018)提供的脑肿瘤MR图像数据集上进行验证,在肿瘤整体区域、肿瘤核心区域和肿瘤增强区域的Dice score分别为0.907、0.838和0.819,与其他方法进行对比,较次优方法分别提升了0.9%、1.3%和0.6%。结论本文方法改进了传统U-Net网络提取和利用图像语义特征不足的问题,并引入了3维MR图像的空间信息,使得肿瘤分割结果更加准确,具有良好的研究和应用价值。  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
9  
被引频次合计:  
12  
ISSN:  
1006-8961  
入藏号:  
CSCD:7068576  
Indexed Date:  
2021-12-25  
Institution Address:  
南京林业大学信息科学技术学院, 南京, 江苏 210037, 中国  
Institution Address:  
浙江树人大学信息科技学院, 杭州, 浙江 310015, 中国  
Institution Address:  
College of Information Science and Technology,Nanjing Forestry University, Nanjing, Jiangsu 210037, China  
Institution Address:  
College of Information Technology,Zhejiang Shuren University, Hangzhou, Zhejiang 310015, China  
Institution:  
南京林业大学信息科学技术学院  
Institution:  
浙江树人大学信息科技学院  
Institution:  
College of Information Science and Technology,Nanjing Forestry University  
Institution:  
College of Information Technology,Zhejiang Shuren University

### 520. Hybrid Multiscale SAR Ship Detector With CNN-Transformer and Adaptive

摘要: Ship detection in remote sensing imagery is crucial for various maritime applications such as surveillance and navigation. Convolutional neural networks (CNNs) and transformers have shown significant potential in object detection within the field of image processing. However, existing models applied directly to ship detection in synthetic aperture radar (SAR) imagery encounter challenges due to the varying sizes of ship targets. This often leads to issues such as low detection accuracy, missed detections, and false alarms. In this letter, we propose a new detection network, HMA-Net, to further address these issues. Initially, we introduce the Cwin module, which enhances interference resistance at a relatively low cost, enabling the model to more accurately capture target information. Subsequently, we design a multiscale ship feature extraction module, which uses a parallel multibranch structure to extract features of ships of various sizes and shapes. Finally, we introduce an adaptive fusion loss function that flexibly allocates loss calculation methods to detected targets, thereby enhancing the robustness of the model and achieving high-quality detection boxes. The proposed HMA-Net achieved improvements of 2.0% and 0.9% in mAP(.50:.95) over the baseline models on the SAR Ship Detection dataset and the High-Resolution SAR Images dataset, using only 3.52 M parameters.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001308225500014  
Indexed Date:  
2024-09-18  
Institution Address:  
Shanghai Maritime Univ, Digital Imaging & Intelligent Comp Lab, Shanghai 201306, Peoples R China  
Institution:

### 521. Faster-RCNN water-floating garbage recognition based on multi-scale

摘要: Aiming at the problems of variable morphology, low resolution and limited information of small-target waterfloating garbage, which lead to unsatisfactory detection results, an improved Faster-RCNN (Faster Regions with Convolutional Neural Network)water-floating garbage detection algorithm was proposed, namely MP-Faster-RCNN (Faster- RCNN with Multi-scale feature and Polarized self-attention). Firstly, a small-target water-floating garbage dataset in Lanzhou part of the Yellow River was established, the combination of atrous convolution and ResNet-50 was used as the backbone feature extraction network instead of the original VGG-16 (Visual Geometry Group 16)to expand the perception field for extracting more small-target features. Secondly, two layers of convolutions of 3\*3 and 1\*1 were set in the Region Proposal Network (RPN)by using multi-scale features to compensate for the feature loss caused by a single sliding window. Finally, polarized self-attention was added before RPN to further utilize multi-scale and channel features to extract finergrained multi-scale spatial information and inter-channel dependencies to generate a feature map with global features, achieving more accurate target box localization. Experimental results show that compared with the original Faster-RCNN, MP-Faster-RCNN can effectively improve the detection accuracy of water-floating garbage with a mean Average Precision (mAP)improvement of 6.37 percentage points, the model size is reduced from 521 MB to 108 MB, and the convergence speed is faster under the same training epoch.  
摘要:  
针对小目标水漂垃圾形态多变、分辨率低且信息有限,导致检测效果不理想的问题,提出一种改进的Faster-RCNN(Faster Regions with Convolutional Neural Network)水漂垃圾检测算法MP-Faster-RCNN(Faster-RCNN with Multi-scale feature and Polarized self-attention)。首先,建立黄河兰州段小目标水漂垃圾数据集,将空洞卷积结合ResNet-50代替原来的VGG-16(Visual Geometry Group 16)作为主干特征提取网络,扩大感受野以提取更多小目标特征;其次,在区域生成网络(RPN)利用多尺度特征,设置3\*3和1\*1的两层卷积,补偿单一滑动窗口造成的特征丢失;最后,在RPN前加入极化自注意力,进一步利用多尺度和通道特征提取更细粒度的多尺度空间信息和通道间依赖关系,生成具有全局特征的特征图,实现更精确的目标框定位。实验结果表明,MP-Faster-RCNN能有效提高水漂垃圾检测精度,与原始Faster-RCNN相比,平均精度均值(mAP)提高了6.37个百分点,模型大小从521 MB降到了108 MB,且在同一训练批次下收敛更快。  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1001-9081  
入藏号:  
CSCD:7686837  
Indexed Date:  
2024-06-14  
Institution Address:  
兰州交通大学电子与信息工程学院, 兰州, 甘肃 730070, 中国  
Institution Address:  
School of Electronics and Information Engineering,Lanzhou Jiaotong University, Lanzhou, Gansu 730070, China  
Institution:  
兰州交通大学电子与信息工程学院  
Institution:  
School of Electronics and Information Engineering,Lanzhou Jiaotong University

### 522. RSANet: Towards Real-Time Object Detection with Residual Semantic-Guided

摘要: The huge computational overhead limits the inference of convolutional neural networks on mobile devices for object detection, which plays a critical role in many real-world scenes, such as face identification, autonomous driving, and video surveillance. To solve this problem, this paper introduces a lightweight convolutional neural network, called RSANet: Towards Real-time Object Detection with Residual Semantic-guided Attention Feature Pyramid Network. Our RSANet consists of two parts: (a) Lightweight Convolutional Network (LCNet) as backbone, and (b) Residual Semantic-guided Attention Feature Pyramid Network (RSAFPN) as detection head. In the LCNet, in contrast to recent advances of lightweight networks that prefer to utilize pointwise convolution for changing the number of feature maps, we design a Constant Channel Module (CCM) to save the Memory Access Cost (MAC) and design Down Sampling Module (DSM) to save the computational cost. In the RSAFPN, meanwhile, we employ Residual Semantic-guided Attention Mechanism (RSAM) to fuse the multi-scale features from LCNet for improving detection performance efficiently. The experiment results show that, on PASCAL VOC 20007 dataset, RSANet only requires 3.24 M model size and needs only 3.54B FLOPs with a 416x416 input image. Compared to YOLO Nano, our method obtains a 6.7% improvement in accuracy and requires less computation. On MS COCO dataset, RSANet only requires 4.35 M model size and needs only 2.34B FLOPs with a 320x320 input image. Our method obtains a 1.3% improvement in accuracy compared to Pelee. The comprehensive experiment results demonstrate that our model achieves promising results in terms of available speed and accuracy trade-off.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
17  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
17  
ISSN:  
1383-469X  
eISSN:  
1572-8153  
入藏号:  
WOS:000604483900001  
Indexed Date:  
2021-01-20  
Institution Address:  
Nanjing Univ Posts & Telecommun, Coll Telecommun & Informat Engn, Nanjing, Peoples R China  
Institution Address:  
Guizhou Normal Univ, Sch Big Data & Comp Sci, Guiyang, Peoples R China  
Institution Address:  
Beijing Elect Sci & Technol Inst, Dept Comp Sci & Technol, Beijing, Peoples R China  
Institution:  
Institution:  
Institution:

### 523. Bi-DAINet: Bi-Directional Discard-Accept-Integrate Network for salient

摘要: Recent deep learning-based salient object detection methods have achieved impressive performance by leveraging multi-level convolutional features, which contain informative and complementary saliency cues. However, it remains a major challenge to make full use of multi-scale feature maps and alleviate the influence of redundant information. To address this problem, we propose a Bi-directional Discard Accept-Integrate Network (Bi-DAINet) for salient object detection. A Discard-Accept-Integrate (DAI) module is exploited to extract meaningful messages from the feature maps and pass it to the subsequent layer. The DAI module discards redundant information of the previous layer, accepts useful information of the current layer, and integrates the remained information from multiple scales. Therefore, noisy responses are suppressed in this process. A bi-directional Convolutional Neural Network (CNN) is designed to enforce messages flowing from low-to-high level and high-to-low level, respectively. We infer a foreground map and a background map by the bi-directional CNN and then fuse them to predict a final saliency map. On one hand, low-level detailed information and high-level semantic feature are potentially integrated; on the other hand, fine-grained saliency cues around the boundary of target objects are explored. An edge-preserving loss function is further employed to ensure the boundary accuracy in prediction results. Extensive experiments on five benchmark datasets demonstrate the effectiveness and robustness of the proposed method. (c) 2020 Elsevier B.V. All rights reserved.  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000611057000013  
Indexed Date:  
2021-02-26  
Institution Address:  
Dalian Univ Technol, Sch Comp Sci & Technol, Dalian 116024, Peoples R China  
Institution Address:  
Dalian Univ Technol, Minist Educ, Key Lab Intelligent Control & Optimizat Ind Equip, Dalian 116024, Peoples R China  
Institution Address:  
Dalian Univ Technol, Sch Innovat & Entrepreneurship, Dalian 116024, Peoples R China  
Institution Address:  
Dalian Univ Technol, Sch Math Sci, Dalian 116024, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 524. EEG classification model for virtual reality motion sickness based on

摘要: Background: Virtual reality motion sickness (VRMS) is a key issue hindering the development of virtual reality technology, and accurate detection of its occurrence is the first prerequisite for solving the issue. Objective: In this paper, a convolutional neural network (CNN) EEG detection model based on multi-scale feature correlation is proposed for detecting VRMS. Methods: The model uses multi-scale 1D convolutional layers to extract multi-scale temporal features from the multi-lead EEG data, and then calculates the feature correlations of the extracted multi-scale features among all the leads to form the feature adjacent matrixes, which converts the time-domain features to correlation-based brain network features, thus strengthen the feature representation. Finally, the correlation features of each layer are fused. The fused features are then fed into the channel attention module to filter the channels and classify them using a fully connected network. Finally, we recruit subjects to experience 6 different modes of virtual roller coaster scenes, and collect resting EEG data before and after the task to verify the model. Results: The results show that the accuracy, precision, recall and F1-score of this model for the recognition of VRMS are 98.66 %, 98.65 %, 98.68 %, and 98.66 %, respectively. The proposed model outperforms the current classic and advanced EEG recognition models. Significance: It shows that this model can be used for the recognition of VRMS based on the resting state EEG.  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0169-2607  
eISSN:  
1872-7565  
入藏号:  
WOS:001240109700001  
PubMed ID:  
38728828  
Indexed Date:  
2024-06-10  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Automat, C IMER, CICAEET, Nanjing 210044, Peoples R China  
Institution Address:  
Yanshan Univ, Dept Elect Engn, Measurement Technol & Instrumentat Key Lab Hebei P, Qinhuangdao 066000, Peoples R China  
Institution:  
Institution:

### 525. Remote Sensing Image Denoising Algorithm with Multi-receptive Field

摘要: Optical remote sensing images contain complex texture features. The noise in remote sensing images affects not only the visual effect of images but also the processing, analysis, transmission, and storage of images. Therefore, image denoising becomes an important step in remote sensing image processing. Traditional denoising methods are likely to cause problems such as loss of image details and blurred denoising results. Recently, deep learning has been rapidly developing in the field of image denoising, compared with traditional algorithms, the stability of the denoising algorithm of deep learning algorithms has improved tremendously. However, the real noise in remote sensing images and the reconstruction of the image after denoising, is the main problem in the field of image denoising at present. In this paper, an MRFENet remote sensing image denoising algorithm based on multi-sensory field feature fusion and enhancement is proposed. To address the problem that image details are lost after denoising and real noise is difficult to be eliminated, the following approach was used. First, a global feature extraction module is introduced, which consists of several convolutions with different dilation rates, followed by fusion of the extracted features. The purpose of this process is to allow the model to expand the receptive field without increasing the number of parameters, and to enable the model to converge quickly by extracting shallow features at different scales. Second, multi-scale feature enhancement blocks are introduced. Each block consists of a multi-scale feature extraction layer and a channel attention module, both of which form the residual structure. The purpose is to be able to extract multi-scale features at different levels and to assign higher weights to important features to achieve enhancement of important features. The residual structure ensures that the network does not explode in gradient due to excessive depth. Finally, in order to reduce the loss of feature information and the fluctuation caused by the fusion of shallow features with deep features, the resulting features at different levels are chosen to be fused step by step to enhance the continuity of pixels. To make the denoised images more consistent with the visual perception, MS-SSIM is chosen as the loss function during the training process. The number of channels and the number of multi-scale feature enhancement blocks of the MRFENet are configurable, and the performance of the network does not increase with the number of modules, so the most suitable network parameters can be obtained by combining the network performance with the computational effort. In order to test the denoising ability of MRFENet for remote sensing images of different sizes, two publicly available remote sensing image datasets with different sizes are selected. By adding different intensity of noise on each dataset, this paper tests the denoising stability of MRFENet for different intensity of noise. In order to test the denoising performance of MRFENet for real noise, a hyperspectral real remote sensing image is selected for testing. PSNR and SSIM are selected as quantitative evaluation metrics for different intensity noise datasets to evaluate the denoising results. NIQE, BRISQUE, PIQE are selected as quantitative no- reference evaluation metrics for real noise datasets to evaluate the denoising results.  
After comparing the denoising results with those of traditional denoising algorithms NLM, BM3D and deep learning algorithms DnCNN, RIDNet and REDJ, it can be concluded that the proposed algorithm has the best performance on each dataset and outperforms other algorithms in all metrics. The images denoised with MRFENet can retain the edge details and do not show excessive smoothing, which is more in line with the visual perception. The effectiveness and generalization of MRFENet algorithm for remote sensing image denoising are verified.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1004-4213  
入藏号:  
WOS:000908417500014  
Indexed Date:  
2023-02-04  
Institution Address:  
Nanjing Univ Sci & Technol, State Key Lab Transient Phys, Nanjing 210094, Peoples R China  
Institution:

### 526. A rotated object detection strategy for remote sensing images using

摘要: Remote sensing image object detection is a challenging task with a focus on detection accuracy. In the object detection task of remote sensing images, scenes often contain objects with varying scales and dense arrays, leading to issues such as low accuracy, missed detections, and false alarms. To address these challenges, this paper introduces MCSC-Net, a rotated object detector, designed to enhance detection performance for objects of different scales in densely arranged scenarios. The proposed detector features a novel neck network named MCFN-V3, which incorporated a cross-fusion structure. MCFN-V3 facilitates simultaneous feature extraction in upper and lower layers with feedback to the middle layer. Through the cross-fusion of feature layers, it achieves the integration of multi-scale feature information, thereby enhancing the network's detection capabilities for objects at multiple scales. Additionally, the paper presents the STC module to address the problem of poor feature correlation in high-resolution image processing by CNN networks. This module improves feature expressiveness and range, strengthening inter-feature relationships within layers. Furthermore, in the target positioning stage, the paper transforms the angle regression problem into a classification task, making the network more suitable for detecting rotated objects in dense scenes. To assess the effectiveness of the algorithm, experiments were conducted on two publicly available remote sensing datasets, DOTA and UCAS-AOD. Our method achieved a multiclass average precision (mAP) of 77.4 % on DOTA and an outstanding performance of 97.0 % on UCASAOD. These experimental results validate the effectiveness of the proposed approach.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1568-4946  
eISSN:  
1872-9681  
入藏号:  
WOS:001308757900001  
Indexed Date:  
2024-09-14  
Institution Address:  
Heilongjiang Univ, Coll Elect Engn, Harbin 150080, Peoples R China  
Institution:

### 527. An Enhanced Lightweight Network for Road Damage Detection Based on Deep

摘要: Achieving accurate and efficient detection of road damage in complex scenes has always been a challenging task. In this paper, an enhanced lightweight network, E-EfficientDet, is proposed. Firstly, a feature extraction enhancement module (FEEM) is designed to increase the receptive field and improve the feature expression capability of the network, which can extract richer multi-scale feature information. Secondly, to promote the reuse of feature information between different layers in the network and take full advantage of multi-scale context information, four pyramid modules with different structures are designed based on the idea of semi-dense connection, among which the bidirectional feature pyramid network with longitudinal connection (LC-BiFPN) is more suitable for road damage detection. Finally, to meet the road damage detection tasks under different hardware resource constraints, the E-EfficientDet-D0 similar to D2 networks are proposed in this paper based on the compound scaling strategy. Experimental results show that the detection accuracy of E-EfficientDet-D0 improves by 2.41% compared with the original EfficientDet-D0 on the publicly available road damage dataset and outperforms other networks such as YOLOv5s, YOLOv7-tiny, YOLOv4-tiny, Faster R-CNN, and SSD. Meanwhile, the detection speed of EfficientDet-D0 can reach 27.0 FPS, which meets the demand for real-time detection, and the model size is only 32.31 MB, which is suitable for deployment in mobile devices such as unmanned inspection carts, UAVs, and smartphones. In addition, the detection accuracy of E-EfficientDet-D2 can reach 57.51%, which is 4.39% higher than E-EfficientDet-D0, and the model size is 61.78 MB, which is suitable for practical application scenarios that require higher detection accuracy and better hardware performance.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
被引频次合计:  
7  
ISSN:  
2079-9292  
入藏号:  
WOS:001014273000001  
Indexed Date:  
2023-07-23  
Institution Address:  
East China Jiaotong Univ, Sch Informat Engn, Nanchang 330013, Peoples R China  
Institution:

### 528. Feature pyramid network algorithm based on context information and

摘要: Aiming at the problem that the classification and localization sub-tasks in object detection require large receptive field and high resolution respectively, and it is difficult to achieve a balance between these two contradictory requirements, a feature pyramid network algorithm based on attention mechanism for object detection was proposed. In the algorithm, multiple different receptive fields were integrated to obtain richer semantic information, multi-scale feature maps were fused in the way of paying more attention to the importance of different feature maps, and the fused feature maps were further refined under the guidance of the attention mechanism. Firstly, multi-scale receptive fields were obtained through multiple atrous convolutions with different dilation rates, which enhanced the semantic information with the preservation of the resolution. Secondly, through the Multi-Level Fusion (MLF), multiple feature maps of different scales were fused after changing to the same resolution through upsampling or pooling operations. Finally, the proposed Attention-guided Feature Refinement Module (AFRM) was used to refine the fused feature maps to enhance semantic information and eliminate the aliasing effect caused by fusion. After replacing the Feature Pyramid Network (FPN) in Faster R-CNN with the proposed feature pyramid, experiments were performed on MS COCO\_2017 dataset. The results show that when the backbone network is ResNet (Residual Network) with a depth of 50 and 101, with the use of the proposed algorithm, the Average Precision (AP) of the model reaches 39.2% and 41.0% respectively, which is 1.4 and 1.0 percentage points higher than that of Faster R-CNN using the original FPN, respectively. It can be seen that the proposed feature pyramid network algorithm can replace the original feature pyramid to be better applied in the object detection scenarios.  
摘要:  
针对目标检测中分类和定位子任务分别需要大感受野和高分辨率,难以在这两个相互矛盾的需求间取得平衡的问题,提出一种用于目标检测的基于注意力机制的特征金字塔网络算法。该算法能整合多个不同感受野来获取更丰富的语义信息,以一种更关注不同特征图重要性的方式融合多尺度特征图,并在注意力机制引导下进一步精练复杂融合后的特征图。首先,通过多尺度的空洞卷积获取多尺度感受野,在保留分辨率的同时增强语义信息;其次,通过多级特征融合(MLF)方式将多个不同尺度的特征图通过上采样或池化操作变为相同分辨率后融合;最后,利用注意力引导的特征精练模块(AFRM)对融合后的特征图作精练处理,丰富语义信息并消除融合带来的混叠效应。将所提特征金字塔替换Faster R-CNN中的特征金字塔网络(FPN)后在MS COCO\_2017数据集上进行实验,结果表明当骨干网络为深度50和101的残差网络(ResNet)时,平均精度(AP)分别达到了39.2%和41.0%,与使用原FPN的Faster R-CNN相比,分别提高了1.4和1.0个百分点。可见,所提特征金字塔网络算法能替代原FPN,更好地应用在目标检测场景中。  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1001-9081  
入藏号:  
CSCD:7570311  
Indexed Date:  
2023-12-29  
Institution Address:  
四川大学计算机学院, 成都, 四川 610065, 中国  
Institution Address:  
College of Computer Science, Sichuan University, Chengdu, Sichuan 610065, China  
Institution:  
四川大学计算机学院  
Institution:  
College of Computer Science, Sichuan University

### 529. Dynamic Q&A multi-label classification based on adaptive multi-scale

摘要: In community question answering (CQA), questioners use labels for question and answer (Q&A) classification when asking questions. Since the answerers do not have the same understanding and perspective of the question, the original labels cannot accurately reflect the Q&A categories with constantly given answers. Therefore, this paper proposes a dynamic Q&A multi-label classification approach based on adaptive multi-scale feature extraction. First, global and local semantic features of Q&As are extracted based on bidirectional long short-term memory network and convolutional neural network models, respectively. Second, the label features extraction and fusion method is proposed. The semantic features of the labels are extracted, the label structure graph based on horizontal and vertical dependencies is constructed, and the label structure and semantic features are fused using the graph attention network integrating multi-head self-attention mechanism. Afterward, the label-aware local features of Q&As are constructed using the attention mechanism and fused with global features of Q&A using the multi-head self-attention, thereby multi-scale fusion classification features of Q&A are established. Then, to adaptively extract the core multi-scale fusion features, a multi-objective feature selection model is established and an improved binary multi-objective Sinh Cosh optimizer algorithm is proposed to solve the model. Finally, a classification prediction layer based on a multilayer perceptron is constructed to obtain the multi-label classification results of Q&A documents. The experimental results based on real Q&A data show the superior performance of the proposed method and validate the effectiveness of the proposed four modules.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1568-4946  
eISSN:  
1872-9681  
入藏号:  
WOS:001406897900001  
Indexed Date:  
2025-02-03  
Institution Address:  
China Univ Petr, Sch Econ & Management, Beijing 102249, Peoples R China  
Institution:

### 530. FSM-YOLO: Apple leaf disease detection network based on adaptive feature

摘要: Apple leaf disease is a key factor affecting apple yield. Detecting apple leaf diseases in unstructured environments presents a significant challenge due to the diverse early forms and varying scales of the diseases, as well as the similarity between the diseased areas and the background. To address these challenges, this paper proposes an improved convolutional neural network FSM-YOLO with adaptive feature capture and spatial context awareness. Firstly, to address the lack of feature extraction due to the complex texture structure of disease features, AFEM (Adaptive Feature Enhancement Module) with the ability of contextual information fusion and channel information modulation is proposed, which enhances the feature extraction capability for multiple disease types. Secondly, SCAA (Spatial Context-aware Attention) module with spatial relationship capture and adaptive receptive field adjustment was designed to enhance the network's ability to spatial relationship modeling and its ability to focus on disease characteristics to distinguish between disease targets and background information. Finally, MKMC (Multi-kernel mixed Convolution) is proposed to enhance multi-scale feature extraction capability by efficiently capturing and integrating information at multiple spatial resolutions to cope with different scales and shape variations of early leaf disease types. Experiments were conducted on an apple leaf disease dataset covering eight different disease types with 15,159 disease instances, and the experimental results show that compared with the baseline model YOLOv8s, FSM-YOLO improves mAP@0.5 by 2.7%, precision by 2.0%, and recall by 4.0%. Meanwhile, experimental results on the open-source apple leaf disease dataset ALDOD and plant leaf disease dataset PlantDoc show that FSM-YOLO outperforms the state-of-the-art algorithms, which validates the versatility of FSM-YOLO and confirms its excellent detection performance in various plant disease scenarios.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
1051-2004  
eISSN:  
1095-4333  
入藏号:  
WOS:001316891400001  
Indexed Date:  
2024-09-27  
Institution Address:  
Northwest Normal Univ, Coll Phys & Elect Engn, Lanzhou 730070, Gansu, Peoples R China  
Institution:

### 531. BiU-net: A dual-branch structure based on two-stage fusion strategy for

摘要: Background and objective: Computer-based biomedical image segmentation plays a crucial role in planning of assisted diagnostics and therapy. However, due to the variable size and irregular shape of the segmentation target, it is still a challenge to construct an effective medical image segmentation structure. Recently, hybrid architectures based on convolutional neural networks (CNNs) and transformers were proposed. However, most current backbones directly replace one or all convolutional layers with transformer blocks, regardless of the semantic gap between features. Thus, how to sufficiently and effectively eliminate the semantic gap as well as combine the global and local information is a critical challenge. Methods: To address the challenge, we propose a novel structure, called BiU-Net, which integrates CNNs and transformers with a two-stage fusion strategy. In the first fusion stage, called Single-Scale Fusion (SSF) stage, the encoding layers of the CNNs and transformers are coupled, with both having the same feature map size. The SSF stage aims to reconstruct local features based on CNNs and long-range information based on transformers in each encoding block. In the second stage, Multi-Scale Fusion (MSF), BiU-Net interacts with multi-scale features from various encoding layers to eliminate the semantic gap between deep and shallow layers. Furthermore, a Context- Aware Block (CAB) is embedded in the bottleneck to reinforce multi-scale features in the decoder. Results: Experiments on four public datasets were conducted. On the BUSI dataset, our BiU-Net achieved 85.50 % on Dice coefficient (Dice), 76.73 % on intersection over union (IoU), and 97.23 % on accuracy (ACC). Compared to the state-of-the-art method, BiU-Net improves Dice by 1.17 %. For the Monuseg dataset, the proposed method attained the highest scores, reaching 80.27 % and 67.22 % for Dice and IoU. The BiU-Net achieves 95.33 % and 81.22 % Dice on the PH2 and DRIVE datasets. Conclusions: The results of our experiments showed that BiU-Net transcends existing state-of-the-art methods on four publicly available biomedical datasets. Due to the powerful multi-scale feature extraction ability, our proposed BiU-Net is a versatile medical image segmentation framework for various types of medical images. The source code is released on (https://github.com/ZYLandy/BiU-Net).  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
1  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0169-2607  
eISSN:  
1872-7565  
入藏号:  
WOS:001292351200001  
PubMed ID:  
38776830  
Indexed Date:  
2024-08-21  
Institution Address:  
Chongqing Univ, Sch Microelect & Commun Engn, Chongqing 400044, Peoples R China  
Institution Address:  
Calif Med Innovat Inst, San Diego, CA 92121 USA  
Institution:  
Institution:  
Calif Med Innovat Inst

## Transformer-based Model

### 1. MMFNet: Multisensor Data and Multiscale Feature Fusion Model for

摘要: Although cross-domain fault diagnosis has received much attention in intelligent mechanical fault diagnosis, most existing methods only achieve knowledge transfer within the data from a single sensor. For a complex industrial system, multiple sensors are usually required to monitor its operating conditions coordinately. In such a situation, the fault diagnosis capability of existing cross-domain methods might be impaired significantly. To address this issue, in this article, a multisensor data and multiscale feature fusion network (MMFNet) is proposed to achieve cross-domain fault diagnosis using multisensor data. More specifically, a multiscale feature extraction module based on the pyramid principle is designed to fuse both deep and shallow features of the complex raw signal. Then, a Transformer-based multiscale features fusion and domain adaptation module is developed to fuse multiscale features and achieve the domain invariant of the sensor data. Finally, the proposed transferable Transformer is used for multisensor features cross-domain fusion. With the developed MMFNet, the features from multiple sensors can be comprehensively captured so that a more accurate fault diagnosis can be achieved. A series of experiments on a complex planetary gearbox demonstrate the effectiveness and practicability of the proposed method for fault diagnosis.

### 2. Sparse Self-Attentive Transformer With Multiscale Feature Fusion on

摘要: The estimation of the state of health (SOH) of lithium-ion batteries (LIBs) plays an important role in ensuring the safe and stable operation of LIB management systems. In order to more accurately predict SOH, a model based on a sparse self-attentive transformer (SSAT) with multitimescale feature fusion is proposed. The SSAT follows an encoder-decoder structure construction, and the model inputs are the extracted health indicators and SOH sequences. The encoder stacks three cross-stage partial (CSP)-ProbSparse attention self-attention blocks, between every two CSP-ProbSparse attention blocks, connections are made by dilated causal convolution and max-pooling layers to obtain exponential growth of the sensory field. All the feature maps output from the self-attention blocks are integrated by multiscale feature fusion, and finally, the appropriate feature dimensions are fed to the decoder through a transition layer to obtain the estimation of SOH. Numerous comparative and ablation experiments have demonstrated that the SSAT model achieves superior performance in a wide range of situations.

### 3. Attention-enhanced multiscale feature fusion network for pancreas and

摘要: Background: Accurate pancreas and pancreatic tumor segmentation from abdominal scans is crucial for diagnosing and treating pancreatic diseases. Automated and reliable segmentation algorithms are highly desirable in both clinical practice and research. Purpose: Segmenting the pancreas and tumors is challenging due to their low contrast, irregular morphologies, and variable anatomical locations. Additionally, the substantial difference in size between the pancreas and small tumors makes this task difficult. This paper proposes an attention-enhanced multiscale feature fusion network (AMFF-Net) to address these issues via 3D attention and multiscale context fusion methods.<br /> Methods: First, to prevent missed segmentation of tumors, we design the residual depthwise attention modules (RDAMs) to extract global features by expanding receptive fields of shallow layers in the encoder. Second, hybrid transformer modules (HTMs) are proposed to model deep semantic features and suppress irrelevant regions while highlighting critical anatomical characteristics. Additionally, the multiscale feature fusion module (MFFM) fuses adjacent top and bottom scale semantic features to address the size imbalance issue.<br /> Results: The proposed AMFF-Net was evaluated on the public MSD dataset, achieving 82.12% DSC for pancreas and 57.00% for tumors. It also demonstrated effective segmentation performance on the NIH and private datasets, outperforming previous State-Of-The-Art (SOTA) methods. Ablation studies verify the effectiveness of RDAMs, HTMs, and MFFM.<br /> Conclusions: We propose an effective deep learning network for pancreas and tumor segmentation from abdominal CT scans. The proposed modules can better leverage global dependencies and semantic information and achieve significantly higher accuracy than the previous SOTA methods.

### 4. M-Swin: Transformer-Based Multiscale Feature Fusion Change Detection

摘要: Remote sensing image change detection is extensively utilized in various applications in the field of remote sensing, particularly in the realm of cropland conservation, where it plays a critical role in protecting the agro-ecosystem and ensuring global food security. However, the progressive improvement in resolution and size of remote sensing imagery has led to a "scale gap" challenge in the detection of small building changes in cropland areas. To address this challenge, an innovative multiscale feature fusion change detection network (M-Swin) based on transformer using hierarchical windows is proposed. In order to obtain clearer edges and better separation of the change results, a novel Siamese transformer encoder (MSW encoder) is proposed, which can better capture the change information in small building through hierarchical windows and fuse the multiscale feature obtained from different windows. To effectively reduce missed and misdetected small area of changing buildings, a novel bitemporal image feature fusion module (BFFM) is proposed, which can enhance the features based on a priori guidance, thus improving the saliency of change regions. Additionally, a new remote sensing image change detection dataset for cropland, called LuojiaSET-CLCD, has been proposed, experimentally demonstrates that M-Swin has good potential for highly accurate change detection of small buildings within cropland areas and outperforms several newly existing methods in three datasets (LEVIR, WHU-CD, and LuojiaSET-CLCD). Our dataset will be publicly available at https://github.com/RSIIPAC/LuojiaSET-CLCD.

### 5. Direction-Guided Multiscale Feature Fusion Network for Geo-Localization

摘要: Cross-view geo-localization has been widely used as an important technique for determining the geographical location of unmanned aerial vehicles (UAVs). Despite various image retrieval methods proposed, drone and satellite image cross-view geo-localization still remain challenging due to their wildly inconsistent view angles. In this article, we propose a new framework, the Swin-radial-locality network (SRLN), to extract robust image feature representations. Specifically, SRLN is based on a pruned version of the Swin transformer, which integrates multiscale feature aggregation within a Siamese network structure, featuring shared weights and equipped with multiclassification heads. SRLN is mainly comprised of a radial-slicer-network (RSN) and a local-pattern-network (LPN), which is designed to effectively harmonize directional information from drone-captured images and broader environmental features from satellite imagery, crucial for capturing angle and feature details between drone and satellite images. The RSN part focuses on capturing fine-grained features that represent the drone's directional information, while the LPN is utilized for a more comprehensive analysis of broader environmental features. Extensive experiments are carried out on widely used public benchmark datasets, i.e., University-1652 and SUES-200. With more than 3% improvement over existing methods in both drone-view target localization tasks and drone navigation applications, the results validate the superior performance of our multiscale feature fusion model, achieving a state-of-the-art performance record.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001226385000040  
Indexed Date:  
2024-06-16  
Institution Address:  
Shanghai Univ Engn Sci, Sch Elect & Elect Engn, Shanghai 201602, Peoples R China  
Institution Address:  
Zhejiang Univ, Jiaxing Res Inst, Intelligent Opt & Photon Res Ctr, Jiaxing 314000, Peoples R China  
Institution:  
Institution:

### 6. Global attention network with multiscale feature fusion for infrared

摘要: A global attention network (GANet) with multiscale feature fusion is proposed to detect infrared small target by introducing a transformer attention module and an adaptive asymmetric fusion module. The transformer attention module is designed to learn the long-range relationship between small targets and background. The adaptive asymmetric fusion module is employed to aggregate the multiscale contextual information from highlevel and low-level features. In addition, a target duplicating data augmentation strategy by copy-pasting small targets many times is proposed to increase the positive samples during training for suppressing the classimbalance problem. Extensive experiments on infrared small target datasets demonstrate that our method can achieve high detection accuracy and low false alarm rate compared with some state-of-the-art model-driven and data-driven methods.  
在中国科学引文数据库中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
22  
被引频次合计:  
24  
ISSN:  
0030-3992  
eISSN:  
1879-2545  
入藏号:  
WOS:001074684600001  
Indexed Date:  
2023-10-08  
Institution Address:  
Cent South Univ, Sch Automat, State Key Lab Precis Mfg Extreme Serv Performance, Changsha 410083, Peoples R China  
Institution:

### 7. A remote sensing image object detection algorithm with improved YOLOv5s

摘要: Aiming at the low average target detection accuracy in remote sensing images caused by obscure features in the objects of interest,complex background information,and multiple small targets,we propose a new remote sensing image object detection algorithm with improved YOLOv5s (Swin-YOLOv5s).First,an efficient channel attention structure is added to the convolutional block of the backbone feature extraction network to suppress the interference of irrelevant information;second,cross-scale connection and contextual information weighting operations are performed to enhance detection target feature extraction on the basis of multiscale feature fusion,and the fused feature maps are composed into a new feature pyramid;finally,the Swin Transformer structure and coordinate attention mechanism are used to further enhance the semantic information and global perception ability of small targets.The result of a feature fusion elimination experiment performed on the DOTA and RSOD datasets shows that the proposed algorithm can significantly improve the average accuracy of object detection in remote sensing images.  
摘要:  
针对遥感图像中感兴趣目标特征不明显、背景信息复杂、小目标居多导致的目标检测精度较低的问题,本文提出了一种改进YOLOv5s的遥感图像目标检测算法(Swin-YOLOv5s)。首先,在骨干特征提取网络的卷积块中加入轻量级通道注意力结构,抑制无关信息的干扰;其次,在多尺度特征融合的基础上进行跨尺度连接和上下文信息加权操作来加强待检测目标的特征提取,将融合后的特征图组成新的特征金字塔;最后,在特征融合的过程中引入Swin Transformer网络结构和坐标注意力机制,进一步增强小目标的语义信息和全局感知能力。将本文提出的算法在DOTA数据集和RSOD数据集上进行消融实验,结果表明,本文提出的算法能够明显提高遥感图像目标检测的平均准确率。  
在中国科学引文数据库中的被引频次:  
7  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
13  
ISSN:  
1673-4785  
入藏号:  
CSCD:7442156  
Indexed Date:  
2023-07-14  
Institution Address:  
华北电力大学(保定)控制与计算机工程学院;;复杂能源系统智能计算教育部工程研究中心, ;;复杂能源系统智能计算教育部工程研究中心, 保定;;保定, 河北;;河北 071003;;071003, 中国  
Institution Address:  
华北电力大学(保定)控制与计算机工程学院, 保定, 河北 071003, 中国  
Institution Address:  
华北电力大学(保定)电气与电子工程学院, 保定, 河北 071003, 中国  
Institution Address:  
School of Control and Computer Engineering,North China Electric Power University(Baoding);;Engineering Research Center of the Ministry of Education for Intelligent Computing of Complex Energy System, ;;Engineering Research Center of the Ministry of Education for Intelligent Computing of Complex Energy System, Baoding;;Baoding, ;; 071003;;071003  
Institution Address:  
School of Control and Computer Engineering,North China Electric Power University(Baoding), Baoding, Hebei 071003, China  
Institution Address:  
School of Electrical and Electronic Engineering,North China Electric Power University(Baoding), Baoding, Hebei 071003, China  
Institution:  
华北电力大学(保定)控制与计算机工程学院;;复杂能源系统智能计算教育部工程研究中心  
Institution:  
华北电力大学(保定)控制与计算机工程学院  
Institution:  
华北电力大学(保定)电气与电子工程学院  
Institution:  
School of Control and Computer Engineering,North China Electric Power University(Baoding);;Engineering Research Center of the Ministry of Education for Intelligent Computing of Complex Energy System  
Institution:  
School of Control and Computer Engineering,North China Electric Power University(Baoding)  
Institution:  
School of Electrical and Electronic Engineering,North China Electric Power University(Baoding)

### 8. Multiscale feature fusion and enhancement in a transformer for the

摘要: Accurate and rapid fine-grained visual classification (FGVC) of tree species within the same family can provide technical support for tree surveys, research, and conservation. However, FGVC faces challenges such as large intraclass differences and small interclass differences. Recognizing tree species within the same family requires focusing on and correlating overall and multiorgan features of the trees while mitigating the influence of complex natural backgrounds, occlusion effects and other factors. To address these challenges, we propose multiscale feature fusion (MFF) and enhancement in transformers to improve recognition performance. The method consists of a Swin transformer backbone, an MFF module, a discriminative feature enhancement (DFE) module, and a texture feature enhancement (TFE) module. The MFF module aims to strike a balance between global and local feature extraction. The DFE module is employed to mitigate the impact of background noise, whereas the TFE module is used to enhance the feature extraction associated with complex textures and spatial patterns. We conducted experiments on a constructed dataset of tree species from the same family, achieving a top-1 accuracy of 90.3 % and a top-3 accuracy of 96.8 %. In addition, the method performed well on three popular FGVC datasets, namely, the Flavia, Oxford Flowers, and PlantCLEF 2015 datasets, with top-1 accuracies of 100 %, 99.2 %, and 81.4 %, respectively. The ablation experiments and module visualizations also yielded satisfactory results. Thus, this work provides a solution to enhance the FGVC task.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1574-9541  
eISSN:  
1878-0512  
入藏号:  
WOS:001405370200001  
Indexed Date:  
2025-02-01  
Institution Address:  
Beijing Forestry Univ, Sch Informat Sci & Technol, Beijing 100083, Peoples R China  
Institution Address:  
Natl Forestry & Grassland Adm, Engn Res Ctr Forestry oriented Intelligent Informa, Beijing 100083, Peoples R China  
Institution:  
Institution:  
Natl Forestry & Grassland Adm

### 9. AMFENet: An Adaptive Multiscale Feature Fusion Enhancement Network for

摘要: Sinkholes are natural depressions or cavities on the Earth's surface, and accurate detection of sinkholes can prevent them from posing significant risks to human life, infrastructure, and the environment. Due to their varied and complex surface manifestations, sinkholes are relatively rare and pose challenges in precise detection when compared with other geological features. To achieve precise sinkhole detection, this letter introduces a network with multiscale feature fusion and enhancement capabilities (AMFENet). AMFENet is the first to propose considering the irregular shapes, varying sizes, relatively small target areas, and uneven distribution characteristics of sinkholes. To capture the contextual global information of sinkholes, AMFENet uses the Swin Transformer as its encoder. To address the scarcity and uneven distribution of sinkholes, a selective amplification block (SAB) is introduced, enhancing the significance of sinkhole features. To handle complex and diverse sinkhole features, the semantic integration block (SIB) is introduced. It integrates both global and local semantic information and adaptively enhances crucial semantic features from both spatial and channel dimensions. According to extensive experiments conducted on the sinkhole DEM dataset in Kentucky, AMFENet outperforms other semantic segmentation methods.  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001174285700001  
Indexed Date:  
2024-04-18  
Institution Address:  
China Univ Geosci, Sch Informat Engn, Beijing 100083, Peoples R China  
Institution:

### 10. FNeXter: A Multi-Scale Feature Fusion Network Based on ConvNeXt and

摘要: The accurate segmentation and quantification of retinal fluid in Optical Coherence Tomography (OCT) images are crucial for the diagnosis and treatment of ophthalmic diseases such as age-related macular degeneration. However, the accurate segmentation of retinal fluid is challenging due to significant variations in the size, position, and shape of fluid, as well as their complex, curved boundaries. To address these challenges, we propose a novel multi-scale feature fusion attention network (FNeXter), based on ConvNeXt and Transformer, for OCT fluid segmentation. In FNeXter, we introduce a novel global multi-scale hybrid encoder module that integrates ConvNeXt, Transformer, and region-aware spatial attention. This module can capture long-range dependencies and non-local similarities while also focusing on local features. Moreover, this module possesses the spatial region-aware capabilities, enabling it to adaptively focus on the lesions regions. Additionally, we propose a novel self-adaptive multi-scale feature fusion attention module to enhance the skip connections between the encoder and the decoder. The inclusion of this module elevates the model's capacity to learn global features and multi-scale contextual information effectively. Finally, we conduct comprehensive experiments to evaluate the performance of the proposed FNeXter. Experimental results demonstrate that our proposed approach outperforms other state-of-the-art methods in the task of fluid segmentation.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
eISSN:  
1424-8220  
入藏号:  
WOS:001210728000001  
PubMed ID:  
38676042  
Indexed Date:  
2024-05-17  
Institution Address:  
Tsinghua Univ, Tsinghua Shenzhen Int Grad Sch, Shenzhen 518055, Peoples R China  
Institution:

### 11. Text detection method combining Segformer with an enhanced feature

摘要: To address the issues of small-scale text omission, text-like pixel misdetection, and inaccurate edge localization in text detection algorithms for natural scenes, we propose a text detection model based on Segformer and an enhanced feature pyramid. First, the model employs an MiT-B2-based encoder to generate multiscale feature maps. Subsequently, during the upsampling phase of the decoder, a cascaded fusion attention module is introduced, which acquires global channel information and text features through global average pooling, global max pooling, and ghost convolution. Then, a two-level orthogonal fusion attention module utilizes asymmetric convolution to enhance the information in the feature fusion section horizontally and vertically. Finally, the results are post-processed using differentiable binarization. The experiments were conducted on the ICDAR2015, ShopSign1265, and MTWI datasets. Compared with the other eight methods, the proposed method achieved the highest F-values, reaching 87.8%, 59.1%, and 74.8%%, respectively. These results demonstrate that the method effectively improves the accuracy of text detection.  
摘要:  
针对自然场景文本检测算法中的小尺度文本漏检、类文本像素误检以及边缘定位不准确的问题,提出一种基于Segformer和增强特征金字塔的文本检测模型。该模型首先采用基于混合Transformer(mix Transformer, MiT)的编码器生成多尺度特征图;然后,在具有特征金字塔结构解码器的上采样部分,提出级联融合注意力模块,通过全局平均池化、全局最大池化和Ghost模块获取全局通道信息并保留文本特征;接着,在解码器的特征融合部分提出两级正交融合注意力模块,利用非对称卷积分别从水平和垂直方向进行信息增强;最后,利用可微分二值化对结果进行后处理。将本文方法在ICDAR2015、ShopSign1265和MTWI 3个数据集上进行实验,相比于其他8种方法,本文方法的F值均为最优,分别达到了87.8%、59.1%和74.8%。结果表明,本文方法有效提高了文本检测的准确率。  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1673-4785  
入藏号:  
CSCD:7852587  
Indexed Date:  
2025-02-09  
Institution Address:  
(保定)华北电力大学控制与计算机工程学院;;(保定)华北电力大学, ;;复杂能源系统智能计算教育部工程研究中心, 保定;;保定, 河北;;河北 071003;;071003, 中国  
Institution Address:  
School of Control and Computer Engineering,North China Electric Power University(Baoding);;Engineering Research Center of intelligent Computing for Complex Energy Systems Ministry of Education, ;;Engineering Research Center of intelligent Computing for Complex Energy Systems Ministry of Education, Baoding;;Baoding, ;; 071003;;071003  
Institution:  
(保定)华北电力大学控制与计算机工程学院;;(保定)华北电力大学  
Institution:  
School of Control and Computer Engineering,North China Electric Power University(Baoding);;Engineering Research Center of intelligent Computing for Complex Energy Systems Ministry of Education

### 12. Multiscale Feature Fusion Booster Network for Segmentation of Colorectal

摘要: Addressing the challenges posed by colorectal polyp variability and imaging inconsistencies in endoscopic images, we propose the multiscale feature fusion booster network (MFFB-Net), a novel deep learning (DL) framework for the semantic segmentation of colorectal polyps to aid in early colorectal cancer detection. Unlike prior models, such as the pyramid vision transformer-based cascaded attention decoder (PVT-CASCADE) and the parallel reverse attention network (PraNet), MFFB-Net enhances segmentation accuracy and efficiency through a unique fusion of multiscale feature extraction in both the encoder and decoder stages, coupled with a booster module for refining fine-grained details and a bottleneck module for efficient feature compression. The network leverages multipath feature extraction with skip connections, capturing both local and global contextual information, and is rigorously evaluated on seven benchmark datasets, including Kvasir, CVC-ClinicDB, CVC-ColonDB, ETIS, CVC-300, BKAI-IGH, and EndoCV2020. MFFB-Net achieves state-of-the-art (SOTA) performance, with Dice scores of 94.38%, 91.92%, 91.21%, 80.34%, 82.67%, 76.92%, and 74.29% on CVC-ClinicDB, Kvasir, CVC-300, ETIS, CVC-ColonDB, EndoCV2020, and BKAI-IGH, respectively, outperforming existing models in segmentation accuracy and computational efficiency. MFFB-Net achieves real-time processing speeds of 26 FPS with only 1.41 million parameters, making it well suited for real-world clinical applications. The results underscore the robustness of MFFB-Net, demonstrating its potential for real-time deployment in computer-aided diagnosis systems and setting a new benchmark for automated polyp segmentation.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0899-9457  
eISSN:  
1098-1098  
入藏号:  
WOS:001447787100001  
Indexed Date:  
2025-03-26  
Institution Address:  
Beijing Univ Technol, Fac Informat Technol, Beijing Key Lab Computat Intelligence & Intelligen, Beijing, Peoples R China  
Institution Address:  
Beijing Univ Technol, Dept Comp Sci, Beijing, Peoples R China  
Institution:  
Institution:

### 13. Radiation Anomaly Detection of Sub-Band Optical Remote Sensing Images

摘要: Radiation anomalies in optical remote sensing images frequently occur due to electronic issues within the image sensor or data transmission errors. These radiation anomalies can be categorized into several types, including CCD, StripeNoise, RandomCode1, RandomCode2, ImageMissing, and Tap. To ensure the retention of image data with minimal radiation issues as much as possible, this paper adopts a self-made radiation dataset and proposes a FlexVisionNet-YOLO network to detect radiation anomalies more accurately. Firstly, RepViT is used as the backbone network with a vision transformer architecture to better capture global and local features. Its multiscale feature fusion mechanism efficiently handles targets of different sizes and shapes, enhancing the detection ability for radiation anomalies. Secondly, a feature depth fusion network is proposed in the Feature Fusion part, which significantly improves the flexibility and accuracy of feature fusion and thus enhances the detection and classification performance of complex remote sensing images. Finally, Inner-CIoU is used in the Head part for edge regression, which significantly improves the localization accuracy by finely adjusting the target edges; Slide-Loss is used for classification loss, which enhances the classification robustness by dynamically adjusting the category probabilities and markedly improves the classification accuracy, especially in the sample imbalance dataset. Experimental results show that, compared to YOLOv8, the proposed FlexVisionNet-YOLO method improves precision, recall, mAP0.5, and mAP0.5:0.9 by 3.5%, 7.1%, 4.4%, and 13.6%, respectively. Its effectiveness in detecting radiation anomalies surpasses that of other models.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001305810700001  
Indexed Date:  
2024-09-10  
Institution Address:  
Minist Nat Resources PR China, Land Satellite Remote Sensing Applicat Ctr, Beijing 100048, Peoples R China  
Institution:

### 14. Joint attention mechanism with dynamic kernel for yolov5 mobile wireless

摘要: The yolov5-CTD (you only look once version five-carafe triplet double attention vision transformer) coil defect detection algorithm is proposed to address the problems associated with the manufacturing process of wireless charging coils, which are the core component of wireless chargers and can produce multiple types of defects. To address the problem of small area defects, a large amount of detail semantic information is lost in the process of down-sampling in the backbone network. DA-ViT (Double Attention Vision transformer) is incorporated into the enhanced feature extraction network to supplement the detail semantic information and enhance the ability of the network to build up the extraction of long-distance information. The Triplet attention mechanism module is introduced to be embedded in the lateral hop connection of the multiscale feature extraction network to enhance the neck network's ability for local information extraction by rotating the dimension for multidimensional feature capture and improve the effect of multiscale feature fusion. For the large number of extreme aspect ratio defects in the coil, Carafe up sampling is used to aggregate contextual information to improve the perceptual field, while optimising the jaggedness and mosaic phenomenon of the defect edges of the feature map caused by nearest neighbor interpolation up sampling. To enable the backbone feature extraction network to learn better feature information, the convolution part of yolov5 is optimised to dynamic conditional convolution. Experiments have shown that yolov5-CTD can detect surface defects on mobile phone wireless charging coils with the accuracy of Map@0.5 reaches 80.9% which is 4.8% higher than the original network, and the detection speed is 41.15FPS, which can meet the industrial production line requirements in terms of speed and accuracy.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1380-7501  
eISSN:  
1573-7721  
入藏号:  
WOS:001022136100006  
Indexed Date:  
2023-08-05  
Institution Address:  
Hohai Univ, Sch Mech & Elect Engn, Changzhou 213022, Peoples R China  
Institution:

### 15. Railway Track Multicomponent Segmentation Based on Residual Contextual

摘要: Regular inspection of track component conditions is crucial for maintaining railway safety, and accurate segmentation of track multicomponent is fundamental to the detection and severity evaluation of track diseases. The U-shaped architecture has achieved tremendous success in segmentation tasks. However, most existing U-shaped architecture-based methods are not suitable for tracking multicomponent segmentation due to the limitations in explicitly modeling long-range dependency. In order to address this problem, this article proposes an innovative track multicomponent segmentation (TMCSeg) network based on the residual contextual transformer (Res-CoT) and adopts the encoder-decoder framework. In the encoder stage, the long-range rich context information can be obtained by adopting designed Res-CoT blocks and cascading transformer layers. In the decoder stage, a multiscale feature fusion (MFF) module is used to aggregate the decoding features of different scales to predict segmentation results. Especially, the decoding features of each scale are obtained by integrating the coding features of the corresponding scale. Numerous tests have been carried out on constructed track datasets. The experimental results demonstrate that our proposed network exhibits strong segmentation capabilities for tracking multicomponent and surpasses the benchmarks in metrics of mIoU and mDice.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001273156700111  
Indexed Date:  
2024-10-16  
Institution Address:  
Hunan Univ, Coll Elect & Informat Engn, Changsha 410082, Peoples R China  
Institution Address:  
Changsha Univ, Coll Elect Informat & Elect Engn, Changsha 410022, Peoples R China  
Institution:  
Institution:

### 16. Human pose estimation in complex background videos via Transformer-based

摘要: Human posture estimation is still a hot research topic. Previous algorithms based on traditional machine learning have difficulties in feature extraction and low fusion efficiency. To address these problems, we proposed a Transformer-based method. We combined three techniques, namely the Transformer-based feature extraction module, the multi-scale feature fusion module, and the occlusion processing mechanism, to capture the human pose. The Transformer-based feature extraction module uses the self-attention mechanism to extract key features from the input sequence, the multi-scale feature fusion module fuses feature information of different scales to enhance the perception ability of the model, and the occlusion processing mechanism can effectively handle occlusion in the data and effectively remove background interference. Our method has shown excellent performance through verification on the standard dataset Human3.6M and the wild video dataset, achieving accurate pose prediction in both complex actions and challenging samples.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
3  
ISSN:  
0141-9382  
eISSN:  
1872-7387  
入藏号:  
WOS:001295085900001  
Indexed Date:  
2024-08-25  
Institution Address:  
Shanghai Univ, Sch Comp Engn & Sci, Shanghai 200444, Peoples R China  
Institution:

### 17. Attention-based multi-scale feature fusion network for myopia grading

摘要: Myopia is a serious threat to eye health and can even cause blindness. It is important to grade myopia and carry out targeted intervention. Nowadays, various studies using deep learning models based on optical coherence tomography (OCT) images to screen for high myopia. However, since regions of interest (ROIs) of pre-myopia and low myopia on OCT images are relatively small, it is rather difficult to use OCT images to conduct detailed myopia grading. There are few studies using OCT images for more detailed myopia grading. To address these problems, we propose a novel attention-based multi-scale feature fusion network named AMFF for myopia grading using OCT images. The proposed AMFF mainly consists of five modules: a pre-trained vision transformer (ViT) module, a multi-scale convolutional module, an attention feature fusion module, an Avg-TopK pooling module and a fully connected (FC) classifier. Firstly, unsupervised pre-training of ViT on the training set can better extract feature maps. Secondly, multi-scale convolutional layers further extract multi-scale feature maps to obtain more receptive fields and extract scale-invariant features. Thirdly, feature maps of different scales are fused through channel attention and spatial attention to further obtain more meaningful features. Lastly, the most prominent features are obtained by the weighted average of the highest activation values of each channel, and then they are used to classify myopia through a fully connected layer. Extensive experiments show that our proposed model has the superior performance compared with other state-of-the-art myopia grading models.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001129743300004  
Indexed Date:  
2024-01-06  
Institution Address:  
Shanghai Jiao Tong Univ, Dept Comupter Sci & Engn, Shanghai, Peoples R China  
Institution Address:  
Shenzhen Univ, Sch Elect & Informat Engn, Shenzhen, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Inst Translat Med, Shanghai, Peoples R China  
Institution Address:  
Huadong Sanat, Wuxi, Jiangsu, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Huadong Sanat

### 18. Multiresolution Aggregation Transformer UNet Based on Multiscale Input

摘要: The latest medical image segmentation methods uses UNet and transformer structures with great success. Multiscale feature fusion is one of the important factors affecting the accuracy of medical image segmentation. Existing transformer-based UNet methods do not comprehensively explore multiscale feature fusion, and there is still much room for improvement. In this paper, we propose a novel multiresolution aggregation transformer UNet (MRA-TUNet) based on multiscale input and coordinate attention for medical image segmentation. It realizes multiresolution aggregation from the following two aspects: (1) On the input side, a multiresolution aggregation module is used to fuse the input image information of different resolutions, which enhances the input features of the network. (2) On the output side, an output feature selection module is used to fuse the output information of different scales to better extract coarse-grained information and fine-grained information. We try to introduce a coordinate attention structure for the first time to further improve the segmentation performance. We compare with state-of-the-art medical image segmentation methods on the automated cardiac diagnosis challenge and the 2018 atrial segmentation challenge. Our method achieved average dice score of 0.911 for right ventricle (RV), 0.890 for myocardium (Myo), 0.961 for left ventricle (LV), and 0.923 for left atrium (LA). The experimental results on two datasets show that our method outperforms eight state-of-the-art medical image segmentation methods in dice score, precision, and recall.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
4  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
15  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
15  
eISSN:  
1424-8220  
入藏号:  
WOS:000801562500001  
PubMed ID:  
35632229  
Indexed Date:  
2022-06-03  
Institution Address:  
Sun Yat Sen Univ, Sch Elect & Commun Engn, Shenzhen 518000, Peoples R China  
Institution:

### 19. NLFFTNet: A non-local feature fusion transformer network for multi-scale

摘要: Recent object detection studies attempt to implement multi-scale feature fusion through complicated hierarchical structures. However, the existing feature fusion methods only focus on the interaction between the same local positions and fail to describe the long-distance dependencies of features. In this study, a novel non-local feature fused transformer convolutional network is proposed for object detection. This model can focus on global semantic information by calculating the attention of different positions to capture the long-distance dependency. Meanwhile, a dynamic data augment method called configurable mix-splicing is introduced to solve the problem of data imbalance between different classes. The experimental results indicate that attributed to the feature fusion and data augment method, our model achieves better performance than state-of-the-art models on two authoritative public datasets. (c) 2022 Elsevier B.V. All rights reserved.  
Web of Science 核心合集中的 "被引频次":  
22  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
23  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:000799997100002  
Indexed Date:  
2022-06-16  
Institution Address:  
Kunming Univ Sci & Technol, Fac Informat Engn & Automat, Kunming 650500, Peoples R China  
Institution Address:  
Kunming Univ Sci & Technol, Yunnan Key Lab Comp Technol Applicat, Kunming 650500, Peoples R China  
Institution Address:  
Tongji Univ, Shanghai Res Inst Intelligent Autonomous Syst, Shanghai 200092, Peoples R China  
Institution:  
Institution:  
Institution:

### 20. Multi-Scale Feature Interactive Fusion Network for RGBT Tracking

摘要: The fusion tracking of RGB and thermal infrared image (RGBT) is paid wide attention to due to their complementary advantages. Currently, most algorithms obtain modality weights through attention mechanisms to integrate multi-modalities information. They do not fully exploit the multi-scale information and ignore the rich contextual information among features, which limits the tracking performance to some extent. To solve this problem, this work proposes a new multi-scale feature interactive fusion network (MSIFNet) for RGBT tracking. Specifically, we use different convolution branches for multi-scale feature extraction and aggregate them through the feature selection module adaptively. At the same time, a Transformer interactive fusion module is proposed to build long-distance dependencies and enhance semantic representation further. Finally, a global feature fusion module is designed to adjust the global information adaptively. Numerous experiments on publicly available GTOT, RGBT234, and LasHeR datasets show that our algorithm outperforms the current mainstream tracking algorithms.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
6  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
eISSN:  
1424-8220  
入藏号:  
WOS:000969067100001  
PubMed ID:  
37050470  
Indexed Date:  
2023-05-04  
Institution Address:  
Sichuan Univ Sci & Engn, Sch Automation & Informat Engn, Yibin 644000, Peoples R China  
Institution Address:  
Sichuan Univ Sci & Engn, Artificial Intelligence Key Lab Sichuan Prov, Yibin 644000, Peoples R China  
Institution:  
Institution:

### 21. Rail-STrans: A Rail Surface Defect Segmentation Method Based on Improved

摘要: With the continuous expansion of the transport network, the safe operation of high-speed railway rails has become a crucial issue. Defect detection on the surface of rails is a key part of ensuring the safe operation of trains. Despite the progress of deep learning techniques in defect detection on the rails' surface, there are still challenges related to various problems, such as small datasets and the varying scales of defects. Based on this, this paper proposes an improved encoder-decoder architecture based on Swin Transformer network, named Rail-STrans, which is specifically designed for intelligent segmentation of high-speed rail surface defects. The problem of a small and black-and-white rail dataset is solved using self-made large and multiple rail surface defect datasets through field shooting, data labelling, and data expansion. In this paper, two Local Perception Modules (LPMs) are added to the encoding network, which helps to obtain local context information and improve the accuracy of detection. Then, the Multiscale Feature Fusion Module (MFFM) is added to the decoding network, which helps to effectively fuse the feature information of defects at different scales in the decoding process and improves the accuracy of defect detection at multiple scales. Meanwhile, the Spatial Detail Extraction Module (SDEM) is added to the decoding network, which helps to retain the spatial detail information in the decoding process and further improves the detection accuracy of small-scale defects. The experimental results show that the mean accuracy of the semantic segmentation of the method proposed in this paper can reach 90.1%, the mean dice coefficient can reach 89.5%, and the segmentation speed can reach 37.83 FPS, which is higher than other networks' segmentation accuracy. And, at the same time, it can achieve higher efficiency.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2076-3417  
入藏号:  
WOS:001220043900001  
Indexed Date:  
2024-05-20  
Institution Address:  
East China Jiaotong Univ, Sch Informat Engn, Nanchang 330013, Peoples R China  
Institution:

### 22. IRFNet: Cognitive-Inspired Iterative Refinement Fusion Network for

摘要: Camouflaged Object Detection (COD) aims to identify objects that are intentionally concealed within their surroundings through appearance, texture, or pattern adaptations. Despite recent advances, extreme object-background similarity causes existing methods struggle with accurately capturing discriminative features and effectively modeling multiscale patterns while preserving fine details. To address these challenges, we propose Iterative Refinement Fusion Network (IRFNet), a novel framework that mimics human visual cognition through progressive feature enhancement and iterative optimization. Our approach incorporates the following: (1) a Hierarchical Feature Enhancement Module (HFEM) coupled with a dynamic channel-spatial attention mechanism, which enriches multiscale feature representations through bilateral and trilateral fusion pathways; and (2) a Context-guided Iterative Optimization Framework (CIOF) that combines transformer-based global context modeling with iterative refinement through dual-branch supervision. Extensive experiments on three challenging benchmark datasets (CAMO, COD10K, and NC4K) demonstrate that IRFNet consistently outperforms fourteen state-of-the-art methods, achieving improvements of 0.9-13.7% across key metrics. Comprehensive ablation studies validate the effectiveness of each proposed component and demonstrate how our iterative refinement strategy enables progressive improvement in detection accuracy.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001443465600001  
PubMed ID:  
40096411  
Indexed Date:  
2025-03-21  
Institution Address:  
Chinese Acad Sci, Shanghai Adv Res Inst, Shanghai 201210, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 100049, Peoples R China  
Institution Address:  
ShanghaiTech Univ, Sch Informat Sci & Technol, Shanghai 201210, Peoples R China  
Institution:  
Institution:  
Institution:

### 23. MFFTNet: A Novel 3D Point Cloud Segmentation Network Based on

摘要: Intelligent analysis of 3D point clouds has become a frontier in emerging fields such as autonomous driving, digital twins, and the metaverse. Precise segmentation of 3D point clouds is particularly important within these domains; however, it faces several challenges: (1) point cloud data inherently lacks structured topological information; (2) point cloud shapes are complex and highly variable, making it difficult to utilize semantic priors; and (3) the sampling process of point clouds may result in sparse and uneven data. To address these issues, this paper proposes a novel Point Cloud Segmentation Network based on multi-scale feature fusion and Transformer architecture (MFFTNet). MFFTNet enhances the performance of existing segmentation methods by globally modeling the overall point cloud shape and embedding local point cloud details. Specifically, MFFTNet divides the segmentation task into encoding and decoding stages. The encoder is designed as a hierarchical pyramid structure that extracts relatively sparse local center points and fuses local features during progressive downsampling. It also utilizes a Transformer for global feature modeling to establish multi-scale topological and semantic information of the point cloud. Subsequently, multi-scale feature fusion further enhances the network's perception of local features and global structure. The decoder progressively upsamples to restore the original point cloud and injects multi-scale feature information to achieve precise segmentation. Based on the aforementioned encoding-decoding structure and multi-scale feature fusion, MFFTNet outperforms existing methods on the point cloud semantic segmentation datasets ShapeNetPart and S3DIS.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2169-3536  
入藏号:  
WOS:001398625000014  
Indexed Date:  
2025-01-23  
Institution Address:  
Hainan Vocat Univ Sci & Technol, Coll Informat Engn, Haikou 571126, Peoples R China  
Institution Address:  
Changzhou Vocat Inst Engn, Sch Intelligent Mfg, Changzhou 213164, Peoples R China  
Institution Address:  
Brisight Hainan Technol & Dev Co Ltd, Haikou 570216, Peoples R China  
Institution Address:  
Hainan Coll Software Technol, Sch Comp & Artificial Intelligence, Qionghai 571400, Peoples R China  
Institution:  
Institution:  
Institution:  
Brisight Hainan Technol & Dev Co Ltd  
Institution:  
Hainan Coll Software Technol

### 24. Small object detection in remote sensing images based on attention

摘要: Due to the influence of dense distribution of detection objects and complex background, there are many small objects, which are difficult to detect in remote sensing images. In order to solve the difficult problem of small object detection in remote sensing images, we propose an object detection algorithm named CotYOLO-v3 in this paper. First, we redesign the residual blocks in the backbone Darknet-53, and we replace it with Contextual Transformer (Cot) blocks with contextual information in the backbone Darknet-53 to extract contextual information for small objects and enhance visual representation; Second, we introduce the shallow information with attention mechanism before the feature fusion of YOLO-v3 to reduce the influence of background interference factors and improve the expression ability of the network. Then, we optimize the feature fusion process, we replace the up-sampling method with sub-pixel convolution, and we replace the first convolution layer of the prediction branch with a residual block. Finally, we use K-Medians clustering algorithm to regenerate the anchors suitable for the remote sensing image datasets. In this paper, we set up a comparative experiment of CotYOLO-v3 and commonly used object detection algorithms to detect small objects in DIOR datasets. The experimental results show that, compared with other commonly used object detection algorithms, CotYOLO-v3 object detection algorithm has obvious advantages in detecting small objects in remote sensing images. Compared with the original object detection algorithm YOLO-v3, the mean Average Precision (mAP) of CotYOLO-v3 improved by 5.07%.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
13  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
13  
ISSN:  
0143-1161  
eISSN:  
1366-5901  
入藏号:  
WOS:000825130200001  
Indexed Date:  
2022-07-21  
Institution Address:  
Yanshan Univ, Hebei Key Lab Measurement Technol & Instrument, Qinhuangdao, Hebei, Peoples R China  
Institution Address:  
Yanshan Univ, Sch Elect Engn, Qinhuangdao, Hebei, Peoples R China  
Institution:  
Institution:

### 25. Attention Guided Multi Scale Feature Fusion Network for Automatic

摘要: The precise and automatic segmentation of prostate magnetic resonance imaging (MRI) images is vital for assisting doctors in diagnosing prostate diseases. In recent years, many advanced methods have been applied to prostate segmentation, but due to the variability caused by prostate diseases, automatic segmentation of the prostate presents significant challenges. In this paper, we propose an attention-guided multi-scale feature fusion network (AGMSFNet) to segment prostate MRI images. We propose an attention mechanism for extracting multi-scale features, and introduce a 3D transformer module to enhance global feature representation by adding it during the transition phase from encoder to decoder. In the decoder stage, a feature fusion module is proposed to obtain global context information. We evaluate our model on MRI images of the prostate acquired from a local hospital. The relative volume difference (RVD) and dice similarity coefficient (DSC) between the results of automatic prostate segmentation and ground truth were 1.21% and 93.68%, respectively. To quantitatively evaluate prostate volume on MRI, which is of significant clinical significance, we propose a unique AGMSF-Net. The essential performance evaluation and validation experiments have demonstrated the effectiveness of our method in automatic prostate segmentation.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001199394600006  
Indexed Date:  
2024-04-24  
Institution Address:  
Hainan Univ, Sch Informat & Commun Engn, Haikou, Peoples R China  
Institution Address:  
Hainan Univ, Coll Comp Sci & Technol, Haikou, Peoples R China  
Institution Address:  
Haikou Municipal Peoples Hosp, Urol Dept, Haikou, Hainan, Peoples R China  
Institution Address:  
Cent South Univ, Affiliated Hosp, Xiangya Med Coll, Haikou, Peoples R China  
Institution Address:  
Hainan Normal Univ, Sch Informat Sci & Technol, Haikou, Peoples R China  
Institution:  
Institution:  
Institution:  
Haikou Municipal Peoples Hosp  
Institution:  
Cent South Univ  
Institution:

### 26. YOLOv5-CSF: an improved deep convolutional neural network for flame

摘要: Fire is a multiple and destructive disaster that usually results in great loss of life and property. Therefore, early detection of fire can help minimize mortality and reduce the risk to ecosystems and property. We propose a novel high-precision flame monitor, YOLOv5-CSF, that can monitor fires using existing monitoring equipment. Based on the YOLOv5 detector, we introduce a coordinate attention mechanism from the backbone network to obtain the relationship between flame position information and channel information in an efficient way which increase the feature expression of the backbone network. The swin transformer block is introduced in the neck network to expand the perceptual field of the network model and improve the flame feature extraction capability. The adaptive spatial feature fusion module is introduced in the head network to strengthen the multi-scale feature fusion of flame features and reduce false alarms. Compared with the original YOLOv5l, the average accuracy of the model is improved by 4.1%. Compared with seven other advanced flame monitors, the proposed algorithm has the highest average accuracy in the flame dataset and the proposed method is quite effective.  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
1432-7643  
eISSN:  
1433-7479  
入藏号:  
WOS:000982753900003  
Indexed Date:  
2023-05-28  
Institution Address:  
Northwest Normal Univ, Coll Phys & Elect Engn, Lanzhou 730000, Peoples R China  
Institution Address:  
Engn Res Ctr Gansu Prov Intelligent Informat Techn, Lanzhou 730000, Peoples R China  
Institution:  
Institution:  
Engn Res Ctr Gansu Prov Intelligent Informat Techn

### 27. MFCTrans: Multi-scale Feature Connection Transformer for Deformable

摘要: Deformable Medical Image Registration (DMIR) aims to establish precise anatomical alignment of multiple medical images. However, the existing U-shape networks encounter difficulties in efficiently transferring multi-scale feature information from the encoder to the decoder. To address this issue, we propose a novel backbone network called MFCTrans, which constructs effective feature connection in DMIR. Drawing inspiration from the attention mechanism observed in the human cognitive system, our proposed method employs a Feature Fusion and Assignment Transformer (FFAT) module and a Spatial Cross Attention Fusion (SCAF) module. The former facilitates the fusion of multi-channel features, while the latter guides the integration of multi-scale information. A Multiple Residual (MR) branch is also deployed between the encoder and FFAT to improve the network's generalization. We conduct extensive qualitative and quantitative evaluations on the OASIS and LPBA40 datasets. The proposed method achieves higher Dice scores than Transmorph by 1.3% and 2.0% on the respective datasets while maintaining a comparable voxel folding percentage. Ablation studies analyze the impacts and efficiency of each component in the proposed method. In summary, our proposed network offers a promising framework for achieving high-quality medical image registration and holds significant potential for applications in computer vision and cognitive computation.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1866-9956  
eISSN:  
1866-9964  
入藏号:  
WOS:001147724100002  
Indexed Date:  
2024-02-02  
Institution Address:  
Shenzhen Univ, Coll Elect & Informat Engn, Shenzhen 518060, Peoples R China  
Institution Address:  
Shenzhen Univ, Guangdong Key Lab Intelligent Informat Proc, Shenzhen 518060, Peoples R China  
Institution:  
Institution:

### 28. DFTI: Dual-Branch Fusion Network Based on Transformer and Inception for

摘要: Due to adverse illumination in space, noncooperative object perception based on multisource image fusion is crucial for on-orbit maintenance and orbital debris removal. In this article, we first propose a dual-branch multiscale feature extraction encoder combining Transformer block (TB) and Inception block (IB) to extract global and local features of visible and infrared images and establish high-dimensional semantic connections. Second, different from the traditional artificial design fusion strategy, we propose a feature fusion module called cross-convolution feature fusion (CCFF) module, which can achieve image feature level fusion. Based on the above, we propose a dual-branch fusion network based on Transformer and Inception (DFTI) for space noncooperative object, which is an image fusion framework based on autoencoder architecture and unsupervised learning. The fusion image can simultaneously retain the color texture details and contour energy information of space noncooperative objects. Finally, we construct a fusion dataset of infrared and visible images for space noncooperative objects (FIV-SNO) and compare DFTI with seven state-of-the-art methods. In addition, object tracking as a follow-up high-level visual task proves the effectiveness of our method. The experimental results demonstrate that compared with other advanced methods, the entropy (EN) and average gradient (AG) of the fusing images using DFTI network are increased by 0.11 and 0.06, respectively. Our method exhibits excellent performance in both quantitative measures and qualitative evaluation.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001239831400017  
Indexed Date:  
2024-06-10  
Institution Address:  
Harbin Inst Technol, Dept Control Sci & Engn, Harbin 150001, Peoples R China  
Institution Address:  
Beijing Inst Spacecraft Syst Engn, Beijing 100094, Peoples R China  
Institution:  
Institution:  
Beijing Inst Spacecraft Syst Engn

### 29. DGA Domain Detection Based on Transformer and Rapid Selective Kernel

摘要: Botnets pose a significant challenge in network security by leveraging Domain Generation Algorithms (DGA) to evade traditional security measures. Extracting DGA domain samples is inherently complex, and the current DGA detection models often struggle to capture domain features effectively when facing limited training data. This limitation results in suboptimal detection performance and an imbalance between model accuracy and complexity. To address these challenges, this paper introduces a novel multi-scale feature fusion model that integrates the Transformer architecture with the Rapid Selective Kernel Network (R-SKNet). The proposed model employs the Transformer's encoder to couple the single-domain character elements with the multiple types of relationships within the global domain block. This paper proposes integrating R-SKNet into DGA detection and developing an efficient channel attention (ECA) module. By enhancing the branch information guidance in the SKNet architecture, the approach achieves adaptive receptive field selection, multi-scale feature capture, and lightweight yet efficient multi-scale convolution. Moreover, the improved Feature Pyramid Network (FPN) architecture, termed EFAM, is utilized to adjust channel weights for outputs at different stages of the backbone network, leading to achieving multi-scale feature fusion. Experimental results demonstrate that, in tasks with limited training samples, the proposed method achieves lower computational complexity and higher detection accuracy compared to mainstream detection models.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001386914000001  
Indexed Date:  
2025-01-05  
Institution Address:  
Chuzhou Univ, Sch Comp & Informat Engn, Chuzhou 239000, Peoples R China  
Institution Address:  
Arizona State Univ, Sch Comp & Augmented Intelligence, Tempe, AZ 85287 USA  
Institution:  
Institution:

### 30. CMFF6D: Cross-modality multiscale feature fusion network for 6D pose

摘要: 6D pose estimation is applied in many fields including robotic grasping tasks. This task becomes particularly challenging in poor lighting conditions or when dealing with texture-less metal objects. Single-modality approaches are insufficient for precise 6D pose estimation. Consequently, it is imperative to exploit the complementary nature of point cloud and RGB information to enhance the accuracy and robustness of 6D pose estimation. To tackle this challenge, we propose a novel framework called CMFF6D, that achieves precise 6D pose estimation by learning the complementary features from the texture information of images, the geometric information of depth images, and the frequency information between the two modalities. Specifically, we propose a new network architecture called the GL-Swin Transformer to extract the image features and use RandLANet to learn the point cloud features. Then, these two kinds of features are adaptively merged with their spatial domain and frequency domain information during the bidirectional fusion process to obtain richer feature representation. In order to improve the pose estimation accuracy of small objects by better capturing the differences between stacked objects, we design a Residual Multiplicative Connection (RMC) block to reduce noise interference in shallow features, thereby obtaining shallow information with more boundary details. Finally, we utilize the Feature Pyramid Aggregation Network for multi-scale feature fusion, achieving a deeper integration of the complementary data. Extensive experiments show that the proposed approach achieves competitive performances on MP6D, LM, and LM-O datasets.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:001406864200001  
Indexed Date:  
2025-02-03  
Institution Address:  
Univ Shanghai Sci & Technol, Sch Mech Engn, Shanghai 200093, Peoples R China  
Institution:

### 31. YOLOv4 with Deformable-Embedding-Transformer Feature Extractor for Exact

摘要: The deep learning method for natural-image object detection tasks has made tremendous progress in recent decades. However, due to multiscale targets, complex backgrounds, and high-scale small targets, methods from the field of natural images frequently fail to produce satisfactory results when applied to aerial images. To address these problems, we proposed the DET-YOLO enhancement based on YOLOv4. Initially, we employed a vision transformer to acquire highly effective global information extraction capabilities. In the transformer, we proposed deformable embedding instead of linear embedding and a full convolution feedforward network (FCFN) instead of a feedforward network in order to reduce the feature loss caused by cutting in the embedding process and improve the spatial feature extraction capability. Second, for improved multiscale feature fusion in the neck, we employed a depth direction separable deformable pyramid module (DSDP) rather than a feature pyramid network. Experiments on the DOTA, RSOD, and UCAS-AOD datasets demonstrated that our method's average accuracy (mAP) values reached 0.728, 0.952, and 0.945, respectively, which were comparable to the existing state-of-the-art methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
eISSN:  
1424-8220  
入藏号:  
WOS:000947408000001  
PubMed ID:  
36904727  
Indexed Date:  
2023-03-30  
Institution Address:  
Cent South Univ Forestry & Technol Univ, Coll Comp & Informat Engn, Changsha 410004, Peoples R China  
Institution:  
Cent South Univ Forestry & Technol Univ

### 32. AE-Net: A High Accuracy and Efficient Network for Railway Obstacle

摘要: The incursion of railway obstacles poses a serious risk to train operations, and numerous accidents occur during train shunting. However, existing algorithms still struggle with finding a compromise between detection accuracy and speed during train movement. Moreover, their accuracy and robustness are inadequate, specifically when handling small objects in complicated railway scenarios. To overcome these issues, this article proposes an efficient network using convolution and transformer (AE-Net) for performing accurate and real-time detection of railway obstacles to ensure driving safety. First, the enhanced and lightweight transformer module (ETM) is constructed to strengthen the model's global modeling ability. Then, the lightweight feature integration module (LIM) is presented to integrate multibranch feature information and reduce model complexity. Finally, the reinforced multiscale feature fusion module (RFM) is utilized to enhance the multiscale object detection capability, especially for small obstacles. The presented algorithm realizes 95.29% mAP and 145 frames/s on the railway dataset, which is superior to YOLOv5s. In addition, the experiment on MS COCO further shows that AE-Net can perform a considerably better detection than current state-of-the-art models. Hence, it is practicable to employ AE-Net in actual railway and further more complex multitarget scenarios.  
Web of Science 核心合集中的 "被引频次":  
3  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001184952300033  
Indexed Date:  
2024-04-12  
Institution Address:  
Tianjin Univ, State Key Lab Precis Measuring Technol & Instrume, Tianjin 300072, Peoples R China  
Institution Address:  
China Univ Min & Technol, Sch Mech Elect & Informat Engn, Beijing 100083, Peoples R China  
Institution:  
Institution:

### 33. UMTF-Net: An Unsupervised Multiscale Transformer Fusion Network for

摘要: Hyperspectral images (HSIs) are extensively utilized in several fields due to their abundant spectral band, particularly for tasks like ground object classification and environmental monitoring. However, as a result of equipment and imaging condition constraints, HSI frequently demonstrates a restricted spatial resolution. The fusion of a low-resolution HSI and a high-resolution multispectral image (HR-MSI) of the same scene is a crucial method for generating an HR-HSI. At present, due to factors, such as complexity and GPU memory limitation, most of the HSI-MSI fusion algorithms based on deep learning (DL) cannot utilize the transformer module well to capture the long-range dependence information in large-size remote sensing images. At the same time, the lack of a large amount of high-quality training data has become an important problem that affects the performance of fusion algorithms based on DL. In response to the above issues, this article introduces a new unsupervised multiscale transformer fusion (UMTF) network, called UMTF-Net, which enables HSI-MSI fusion without the need for additional training data. UMTF-Net is composed of an HSI fusion network and a U-network (U-Net)-based multiscale feature extraction network. In order to learn the cross-feature spatial similarity and long-range dependency of MSI and HSI, we first extract the multiscale features of MSI using the U-Net-based multiscale feature extraction network. We then input these features into the corresponding scale cross-feature fusion transformer module in the HSI fusion network to conduct feature fusion. Then, we input the fused features into the spatial spectral fuse attention module for spatial spectral feature enhancement, and finally generate HR-HSI. Comparing UMTF-Net to other advanced methods, the fusion results from three datasets and multiple ablation experiments indicate that our method performs excellently in different evaluations.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001329047300003  
Indexed Date:  
2024-10-15  
Institution Address:  
Hebei Univ, Coll Elect & Informat Engn, Baoding 071002, Peoples R China  
Institution Address:  
Machine Vis Technol Innovat Ctr Hebei Prov, Baoding 071002, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Automat, State Key Lab Multimodal Artificial Intelligence S, Beijing 100190, Peoples R China  
Institution Address:  
Univ Leicester, Sch Comp & Math Sci, Leicester LE1 7RH, England  
Institution:  
Institution:  
Machine Vis Technol Innovat Ctr Hebei Prov  
Institution:  
Institution:

### 34. An Efficient UAV Image Object Detection Algorithm Based on Global

摘要: Object detection technology holds significant promise in unmanned aerial vehicle (UAV) applications. However, traditional methods face challenges in detecting denser, smaller, and more complex targets within UAV aerial images. To address issues such as target occlusion and dense small objects, this paper proposes a multi-scale object detection algorithm based on YOLOv5s. A novel feature extraction module, DCNCSPELAN4, which combines CSPNet and ELAN, is introduced to enhance the receptive field of feature extraction while maintaining network efficiency. Additionally, a lightweight Vision Transformer module, the CloFormer Block, is integrated to provide the network with a global receptive field. Moreover, the algorithm incorporates a three-scale feature fusion (TFE) module and a scale sequence feature fusion (SSFF) module in the neck network to effectively leverage multi-scale spatial information across different feature maps. To address dense small objects, an additional small object detection head was added to the detection layer. The original large object detection head was removed to reduce computational load. The proposed algorithm has been evaluated through ablation experiments and compared with other state-of-the-art methods on the VisDrone2019 and AU-AIR datasets. The results demonstrate that our algorithm outperforms other baseline methods in terms of both accuracy and speed. Compared to the YOLOv5s baseline model, the enhanced algorithm achieves improvements of 12.4% and 8.4% in AP50 and AP metrics, respectively, with only a marginal parameter increase of 0.3 M. These experiments validate the effectiveness of our algorithm for object detection in drone imagery.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
2079-9292  
入藏号:  
WOS:001341701900001  
Indexed Date:  
2024-11-02  
Institution Address:  
Nanyang Technol Univ, Sch Elect & Elect Engn, Singapore 639798, Singapore  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Coll Automat Engn, Nanjing 210016, Peoples R China  
Institution:  
Institution:

### 35. R-YOLOv5: A Lightweight Rotational Object Detection Algorithm for

摘要: A lightweight rotational object detection algorithm, R-YOLOv5, is proposed to address the limitations of traditional object detection algorithms that do not consider the diversity of vehicle scales in drone images and fail to obtain information on rotation angles. The proposed algorithm incorporated an angle prediction branch and introduced a circular smooth label (CSL) angle classification method to make it suitable for detection scenarios based on rotational boxes. A cascaded Swin Transformer block (STrB) is used to reduce computational complexity during feature fusion in the backbone network, further enhancing semantic information and global perception capabilities for small objects. A feature enhancement attention module (FEAM) is proposed to improve the utilization of detailed information through local feature self-supervision. An adaptive spatial feature fusion structure (ASFF) is introduced, which employs features extracted from different levels of the backbone network to perform multi-scale feature fusion. The experimental results show that the detection accuracy reaches 84.91% on the Drone-Vehicle dataset and 90.23% on the UCAS-AOD remote sensing dataset. The lightweight model has a parameter count of only 2.02 million and can achieve 82.6 FPS for high-resolution images, which is significantly better than existing lightweight models and more suitable for real-time detection of rotating vehicles in dense scenes, making it suitable for deployment on a large majority of embedded platforms.  
Web of Science 核心合集中的 "被引频次":  
12  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
12  
ISSN:  
2169-3536  
入藏号:  
WOS:001018598300001  
Indexed Date:  
2023-07-09  
Institution Address:  
Shanghai Univ Elect Power, Sch Elect & Informat Engn, Shanghai 200090, Peoples R China  
Institution Address:  
Fudan Univ, Acad Engn & Technol, Shanghai 200433, Peoples R China  
Institution:  
Institution:

### 36. PolSAR image classification using complex-valued multiscale attention

摘要: This paper Introduces a novel method for Polarimetric Synthetic Aperture Radar (PolSAR) image classification using a Complex-Valued Multiscale Attention Vision Transformer (CV-MsAtViT). The model incorporates a complex-valued multiscale feature fusion mechanism, a complex-valued attention block, and a Complex-Valued Vision Transformer (CV-ViT) to effectively capture spatial and polarimetric features from PolSAR data. The multiscale fusion block enhances feature extraction, while the attention mechanism prioritizes critical features, and the CV-ViT processes data in the complex domain, preserving both amplitude and phase information. Experimental results on benchmark PolSAR datasets, including Flevoland, San Francisco, and Oberpfaffenhofen, show that CV-MsAtViT achieves superior classification accuracy, with an overall accuracy (OA) of 98.35% on the Flevoland dataset, outperforming state-of-the-art models like PolSARFormer. The model also demonstrates efficient computational performance, minimizing the number of parameters while preserving high accuracy. These results confirm that CV-MsAtViT effectively enhances the classification of PolSAR images by leveraging complex-valued data processing, offering a promising direction for future advancements in remote sensing and complex-valued deep learning. The codes associated with this paper are publicly available at https://github.com/mqalkhatib/CV-MsAtViT.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1569-8432  
eISSN:  
1872-826X  
入藏号:  
WOS:001429274800001  
Indexed Date:  
2025-03-01  
Institution Address:  
Univ Dubai, Coll Engn & IT, Dubai 14143, U Arab Emirates  
Institution:

### 37. A circuit breaker moving contact tracking methods based on convolution

摘要: Measuring the motion characteristics of circuit breaker moving contacts can help diagnose the operating status of the circuit breaker. Currently, most measurement methods are "contact" testing methods, which generally have problems with inconvenient installation and low measurement accuracy. Therefore, a new model that can achieve non-contact measurement method is proposed. Firstly, the multi-scale feature fusion structure is used to fuse the extracted multi-layer depth features. Secondly, the improved Transformer structure with introduced convolution operation is used for feature enhancement. Finally, the prediction head is used to predict the tracking results. Experimental analysis shows that compared with the original algorithm, the tracking success rate of the tracking algorithm has increased by 2.6%, and the precision has increased by 13.9%. The model can achieve accurate tracking and obtain the circuit breaker stroke time curve, which can reasonably reflect the action char-acteristics of the circuit breaker operating mechanism.  
摘要:  
测量断路器动触头运动特性有助于断路器运行状态的故障诊断。目前大部分测量方法为接触式测试,普遍存在安装不便、测量精度低的问题。为此,提出一种新的可以实现非接触式测量的算法。首先利用多尺度特征融合结构融合提取的多层深度特征,其次利用引入卷积操作的改进Transformer结构进行特征增强,最后通过预测头预测跟踪结果。实验结果表明,该跟踪算法相较于原算法,跟踪成功率提升了2.6%,精确度提升了13.9%,可以实现准确跟踪,进而得到断路器行程-时间曲线图,合理地反映了断路器操动机构的动作特性。  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1007-130X  
入藏号:  
CSCD:7547111  
Indexed Date:  
2023-12-08  
Institution Address:  
华北电力大学(保定)计算机系, 保定, 河北 071003, 中国  
Institution Address:  
Department of Computer Science,North China Electric Power University (Baoding), Baoding, Hebei 071003, China  
Institution:  
华北电力大学(保定)计算机系  
Institution:  
Department of Computer Science,North China Electric Power University (Baoding)

### 38. Super-resolution Reconstruction of Remote Sensing Image Based on

摘要: To address the limitation of the existing super-resolution reconstruction of remote sensing image algorithms in fully extracting and utilizing features and coping with high computational complexity in complex scenes, a Transformer network model for super-resolution reconstruction of remote sensing image based on multi-scale feature fusion was proposed. The multi-scale residual Swin Transformer module was introduced to fully extract features and reduce the module redundancy used for flat feature extraction. A feature fusion refinement module was established that can fully extract image features to improve network performance. Based on the public UC Merced Land Use dataset, the experimental results show that the number of parameters required by the proposed model is only 61.6% of the parameters compared with the current mainstream super-resolution reconstruction method EDSR model. The peak signal-to-noise ratio and structural similarity of the reconstruction results at different scales are increased by 0.82 dB and 0.024 on average compared with the EDSR model. Through comparative analysis, it is proved that the model proposed can effectively reduce the redundancy of network parameters while improving the quality of the image. It can significantly improve the quality of the reconstructed image to meet the requirements of high-resolution remote sensing image processing.  
摘要:  
针对现有遥感影像超分辨率重建算法,在处理复杂场景时,存在无法充分提取和利用特征,且计算复杂度高的问题,提出一种多尺度特征融合的Transformer遥感影像超分辨率重建网络模型.该模型引入了多尺度残差Swin Transformer模块,在充分提取特征的同时,减少用于提取浅层特征的模块冗余;建立了一个特征细化融合模块,可以充分提取图像特征来提高网络性能.基于UC Merced Land Use公开数据集进行实验,结果表明:提出的模型所需参数数量仅为目前主流超分辨率重建方法EDSR模型的61.6%,重建结果在不同尺度下的峰值信噪比和结构相似度相对EDSR分别平均提高了0.82 dB和0.024.通过对比分析,证明本文提出的模型在提高图像质量的同时,有效地减少了网络参数冗余,可明显提高重建图像质量,满足高分辨率遥感影像处理需要.  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1005-3026  
入藏号:  
CSCD:7845828  
Indexed Date:  
2025-01-25  
Institution Address:  
东北大学资源与土木工程学院, 沈阳, 辽宁 110819, 中国  
Institution Address:  
School of Resources & Civil Engineering,Northeastern University, Shenyang, Liaoning 110819, China  
Institution:  
东北大学资源与土木工程学院  
Institution:  
School of Resources & Civil Engineering,Northeastern University

### 39. Design of Swin Transformer for semantic segmentation of road scenes

摘要: Road scene semantic segmentation is a crucial task in autonomous driving environment perception. In recent years, Transformer neural networks have been applied in the field of computer vision and have shown excellent performance. Addressing issues such as low semantic segmentation accuracy in complex scene images and insufficient recognition capabilities for small objects, this paper proposes a road scene semantic segmentation algorithm based on Swin Transformer with multiscale feature fusion. The network adopts an encoder-decoder structure, where the encoder utilizes an improved Swin Transformer feature extractor for road scene image feature extraction. The decoder consists of an attention fusion module and a feature pyramid network, effectively integrating semantic features at multiple scales. Validation tests on the Cityscapes urban road scene dataset show that, compared to various existing semantic segmentation algorithms, our approach demonstrates significant improvement in segmentation accuracy.  
摘要:  
道路场景语义分割是自动驾驶环境感知的一项重要任务。近年来,变换神经网络(Transformer)在计算机视觉领域开始应用并取得了很好的效果。针对复杂场景图像语义分割精度低、细小目标识别能力不足等问题,本文提出了一种基于移动窗口Transformer的多尺度特征融合的道路场景语义分割算法。该网络采用编码-解码结构,编码器使用改进后的移动窗口Transformer特征提取器对道路场景图像进行特征提取,解码器由注意力融合模块和特征金字塔网络构成,充分融合多尺度的语义特征。在Cityscapes城市道路场景数据集上进行验证测试,实验结果表明,与多种现有的语义分割算法进行对比,本文方法在分割精度方面有较大的提升。  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1003-501X  
入藏号:  
CSCD:7714868  
Indexed Date:  
2024-07-20  
Institution Address:  
上海理工大学光电信息与计算机工程学院, 上海 200093, 中国  
Institution Address:  
School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China  
Institution:  
上海理工大学光电信息与计算机工程学院  
Institution:  
School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology

### 40. Fast Semantic Segmentation of Ultra-High-Resolution Remote Sensing

摘要: For ultra-high-resolution (UHR) image semantic segmentation, striking a balance between computational efficiency and storage space is a crucial research direction. This paper proposes a Feature Fusion Network (EFFNet) to improve UHR image semantic segmentation performance. EFFNet designs a score map that can be embedded into the network for training purposes, enabling the selection of the most valuable features to reduce storage consumption, accelerate speed, and enhance accuracy. In the fusion stage, we improve upon previous redundant multiple feature fusion methods by utilizing a transformer structure for one-time fusion. Additionally, our combination of the transformer structure and multibranch structure allows it to be employed for feature fusion, significantly improving accuracy while ensuring calculations remain within an acceptable range. We evaluated EFFNet on the ISPRS two-dimensional semantic labeling Vaihingen and Potsdam datasets, demonstrating that its architecture offers an exceptionally effective solution with outstanding semantic segmentation precision and optimized inference speed. EFFNet substantially enhances critical performance metrics such as Intersection over Union (IoU), overall accuracy, and F1-score, highlighting its superiority as an architectural innovation in ultra-high-resolution remote sensing image semantic segmentation.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
eISSN:  
2072-4292  
入藏号:  
WOS:001311563300001  
Indexed Date:  
2024-09-21  
Institution Address:  
Tsinghua Univ, Sch Architecture, Dept Landscape Architecture, Beijing 100084, Peoples R China  
Institution Address:  
China Urban Construct Design & Res Inst, Urban Rural Ecol Landscape Construction & Res Inst, Beijing 100120, Peoples R China  
Institution Address:  
Dalian Univ, Life Sci & Technol Coll, Dept Hort, Dalian 116622, Peoples R China  
Institution Address:  
Beijing Forestry Univ, Sch Art & Design, Dept Environm Art Design, Beijing 100083, Peoples R China  
Institution:  
Institution:  
China Urban Construct Design & Res Inst  
Institution:  
Institution:

### 41. Siamese Tracking Network with Multi-attention Mechanism

摘要: Object trackers based on Siamese networks view tracking as a similarity-matching process. However, the correlation operation operates as a local linear matching process, limiting the tracker's ability to capture the intricate nonlinear relationship between the template and search region branches. Moreover, most trackers don't update the template and often use the first frame of an image as the initial template, which will easily lead to poor tracking performance of the algorithm when facing instances of deformation, scale variation, and occlusion of the tracking target. To this end, we propose a Simases tracking network with a multi-attention mechanism, including a template branch and a search branch. To adapt to changes in target appearance, we integrate dynamic templates and multi-attention mechanisms in the template branch to obtain more effective feature representation by fusing the features of initial templates and dynamic templates. To enhance the robustness of the tracking model, we utilize a multi-attention mechanism in the search branch that shares weights with the template branch to obtain multi-scale feature representation by fusing search region features at different scales. In addition, we design a lightweight and simple feature fusion mechanism, in which the Transformer encoder structure is utilized to fuse the information of the template area and search area, and the dynamic template is updated online based on confidence. Experimental results on publicly tracking datasets show that the proposed method achieves competitive results compared to several state-of-the-art trackers.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1370-4621  
eISSN:  
1573-773X  
入藏号:  
WOS:001296721900001  
Indexed Date:  
2024-08-28  
Institution Address:  
Dalian Minzu Univ, Sch Informat & Commun Engn, Liaohexi Rd, Dalian 116600, Liaoning, Peoples R China  
Institution Address:  
Harbin Inst Technol, Dept Informat Engn, Xidazhi St, Harbin 150006, Heilongjiang, Peoples R China  
Institution:  
Institution:

### 42. Radiology Report Generation Method Based on Multi-scale Feature Parsing

摘要: When using deep learning models to automatically generate radiology reports, due to the extreme imbalance of data, it is difficult for current models to identify abnormal regional features, which leads to misjudgment and missed judgment of the disease. In order to improve the model's ability to identify diseases and improve the quality of reports, the authors use a multi-scale feature parsing Transformer (MFPT) model to generate radiology reports. Among them, a key feature enhanced attention (KFEA) module is constructed to strengthen the utilization of key features. A multi-modal feature fusion (MFF) module is designed to promote the feature fusion of semantic features and visual features and alleviate the impact caused by feature differences. This paper explores the role of stage-aware (SA) module in optimizing primary features in radiology reporting tasks. Finally, compared with the current mainstream models on the popular radiology report dataset IU X-Ray, the results show that the proposed model has achieved the current best effect.  
摘要:  
在使用深度学习模型自动生成放射学报告时,由于数据的极度不平衡,当前的模型难以识别异常区域特征,从而导致对疾病的错判与漏判。为了提升模型对疾病的识别能力,提高放射学报告的质量,提出使用多尺度特征解析Transformer (MFPT)模型来生成放射学报告。构建一个关键特征强化注意力(KFEA)模块,以便加强对关键特征的利用;设计一个多模态特征融合(MFF)模块,以便促进语义特征与视觉特征的特征融合,缓解特征差异造成的影响;探索阶段感知(SA)模块在放射学报告任务中对初级特征的优化作用。最后,在流行的放射学报告数据集IU X-Ray上,与当前的主流模型进行对比实验,结果表明,所提模型取得当前最佳效果。  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0479-8023  
入藏号:  
CSCD:7638899  
Indexed Date:  
2024-04-04  
Institution Address:  
山东科技大学计算机科学与工程学院, 青岛, 山东 266590, 中国  
Institution Address:  
College of Computer Science and Engineering,Shandong University of Science and Technology, Qingdao, Shandong 266590, China  
Institution:  
山东科技大学计算机科学与工程学院  
Institution:  
College of Computer Science and Engineering,Shandong University of Science and Technology

### 43. Semi-DinatNet: Two-Stage Underwater Image Enhancement With

摘要: Underwater image enhancement (UIE) is crucial for marine surveys, attracting significant research interest. The underwater environment often results in images with low illumination, blurriness, and color distortion, challenging the effectiveness of underwater image enhancement. Despite progress in deep learning-based UIE methods, many still struggle with brightness enhancement and detail restoration. Additionally, most methods rely on supervised learning, which performs poorly on unlabeled underwater images. We introduce a two-stage model for enhancing underwater images. The first stage employs a U-shaped Transformer with the dilated neighborhood attention transformer (DNAT) as the bottleneck layer and the channel-wise multi-scale feature fusion transformer (CMSFFT) in the up-sampling and down-sampling phases. DNAT captures long-range dependencies and models global features, Improving the network's capacity to concentrate on both overall image areas and detailed textures. CMSFFT improves attention to attenuated color channels. The second stage includes an illumination-perception processing branch as a subsidiary network. By introducing an illumination guidance block (IGB) between the multi-dilated convolution block (MDB) and residual contextual block (RCB), the network better perceives color and light source information. Feature fusion is achieved through attention feature fusion (AFF), integrating the illumination-perception processing branch with the backbone network. Our training paradigm uses an optimized knowledge distillation approach for semi-supervised learning, enhancing the model's efficiency with unlabeled underwater images and ensuring robust performance across various scenarios. Numerous experiments across various underwater image datasets confirm the exceptional performance of the proposed method.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001340725300001  
Indexed Date:  
2024-10-30  
Institution Address:  
Jiangsu Univ Sci & Technol, Ocean Coll, Zhenjiang 212003, Peoples R China  
Institution Address:  
Jiangsu Univ Sci & Technol, Sch Naval Architecture & Ocean Engn, Zhenjiang 212003, Peoples R China  
Institution:  
Institution:

### 44. An improved multi-scale feature extraction network for medical image

摘要: Background: The use of U-Net and its variations has led to significant advancements in medical image segmentation. However, the encoder-decoder structures of these models often lose spatial information during downsampling. Skip connections can help address this issue; however, they may also introduce excessive irrelevant background information. Additionally, medical images display significant scale variations and complex tissue structures, making it challenging for existing models to accurately separate tissues from the background. To address these issues, we developed the Res2Net-ConvFormer-Dilation-UNet (Res2CD-UNet), a multi-scale feature extraction network for medical image segmentation.<br /> Methods: This study presents a novel U-shaped segmentation network that employs Res2Net as the backbone and incorporates a convolution-style transformer in the encoding stage to enhance global attention. Additionally, a novel channel feature fusion block (CFFB) has been introduced in the skip connection stage to minimize the effects of background noise.<br /> Results: The proposed model was evaluated using publicly available datasets, Synapse and Seg.A.2023. Using the Synapse dataset, the average dice similarity coefficient (DSC) reached 83.92%, which was 1.96% higher than the suboptimal model, and the average Hausdorff distance (HD) was 14.51 mm. Among the eight organs evaluated, optimal results were achieved for four organs. Similarly, using the Seg.A.2023 dataset, the proposed model also achieved the best results with an average DSC of 93.27%.<br /> Conclusions: The results of this study indicate that the proposed model can more accurately segment regions of interest and better extract multi-scale features in medical images than existing deep-learning algorithms.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2223-4292  
eISSN:  
2223-4306  
入藏号:  
WOS:001389784900007  
PubMed ID:  
39698667  
Indexed Date:  
2025-01-09  
Institution Address:  
Univ Shanghai Sci & Technol, Sch Energy & Power Engn, 516 Jungong Rd, Shanghai 200093, Peoples R China  
Institution Address:  
Shanghai Key Lab Multiphase Flow & Heat Transfer P, 516 Jungong Rd, Shanghai 200093, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Key Lab Power Machinery & Engn, Minist Educ, 800 Dongchuan Rd, Shanghai 200240, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Inst Pediat Translat Med, Shanghai Childrens Med Ctr, Sch Med, Shanghai, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Shanghai Engn Res Ctr Virtual Real Struct Heart Di, Shanghai Childrens Med Ctr, Sch Med, Shanghai, Peoples R China  
Institution Address:  
Shanghai Jiao Tong Univ, Shanghai Inst Pediat Congenital Heart Dis, Shanghai Childrens Med Ctr, Sch Med, Shanghai, Peoples R China  
Institution:  
Institution:  
Shanghai Key Lab Multiphase Flow & Heat Transfer P  
Institution:  
Institution:  
Institution:  
Institution:

### 45. Object Detection for Remote Sensing Based on the Enhanced YOLOv8 With

摘要: To address the challenges of object detection in complex remote sensing imagery, where the YOLO backbone network struggles with adaptive learning of feature distributions, leading to insufficient multi-scale feature learning capabilities and low detection accuracy for small and occluded objects, the lightweight Enhanced YOLOv8 with WBiFPN (Weighted Bidirectional Feature Pyramid Network) model is introduced in this paper. This model is designed to enhance multi-scale feature learning performance. It incorporates a feature fusion network based on WBiFPN and introduces the EMA (Efficient Multi-Scale Attention Module) to strengthen the representation of semantic and spatial information, thereby deepening the integration of multi-scale features. The model integrates RepConv (Re-parameterized Convolution) and ConvNeXt C2f in the shallow layers of the backbone network to optimize feature extraction, while the deeper layers include a BoT (Bottleneck Transformer Model) to further enhance multi-scale feature extraction capabilities. To reduce model parameters and computational complexity, the neck network employs a simplified Slim-Neck structure. Experimental results demonstrate that the Enhanced YOLOv8 model exhibits superior performance. Compared to the YOLOv8-n/s/m/l/x series models, the proposed model achieves mean Average Precision (mAP@0.5) of 94.8%, 91.6%, and 82.0% on the NWPU VHR-10, DIOR, and DOTA datasets, respectively, representing improvements of 3.2%, 2.5%, and 2.5%. The average inference speeds are 82 fps, 79 fps, and 76 fps, meeting the real-time requirements of inference. Furthermore, the Enhanced YOLOv8 model outperforms other mainstream models in detection performance.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
2169-3536  
入藏号:  
WOS:001346700600001  
Indexed Date:  
2024-11-08  
Institution Address:  
Taiyuan Inst Technol, Dept Elect Engn, Taiyuan 030008, Peoples R China  
Institution Address:  
Changchun Univ Sci & Technol, Sch Elect & Informat Engn, Changchun 130022, Peoples R China  
Institution Address:  
Changchun Univ Sci & Technol, Overseas Expertise Intro Project Discipline Innova, Changchun 130022, Peoples R China  
Institution:  
Institution:  
Institution:

### 46. Micro-Gear Point Cloud Segmentation Based on Multi-Scale Point

摘要: To address the challenges in industrial precision component detection posed by existing point cloud datasets, this research endeavors to amass and construct a point cloud dataset comprising 1101 models of miniature gears. The data collection and processing procedures are elaborated upon in detail. In response to the segmentation issues encountered in point clouds of small industrial components, a novel Point Transformer network incorporating a multiscale feature fusion strategy is proposed. This network extends the original Point Transformer architecture by integrating multiple global feature extraction modules and employing an upsampling module for contextual information fusion, thereby enhancing its modeling capabilities for intricate point cloud structures. The network is trained and tested on the self-constructed gear dataset, yielding promising results. Comparative analysis with the baseline Point Transformer network indicates a notable improvement of 1.1% in mean Intersection over Union (mIoU), substantiating the efficacy of the proposed approach. To further assess the method's effectiveness, several ablation experiments are designed, demonstrating that the introduced modules contribute to varying degrees of segmentation accuracy enhancement. Additionally, a comparative evaluation is conducted against various state-of-the-art point cloud segmentation networks, revealing the superior performance of the proposed methodology. This research not only aids in quality control, structural detection, and optimization of precision industrial components but also provides a scalable network architecture design paradigm for related point cloud processing tasks.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
eISSN:  
2076-3417  
入藏号:  
WOS:001232938600001  
Indexed Date:  
2024-06-03  
Institution Address:  
Beijing Union Univ, Beijing Key Lab Informat Serv Engn, Beijing 100101, Peoples R China  
Institution Address:  
Beijing Union Univ, Coll Robot, Beijing 100101, Peoples R China  
Institution Address:  
Beijing Union Univ, Beijing Intelligent Machinery Innovat Design Serv, Beijing 100101, Peoples R China  
Institution:  
Institution:  
Institution:

### 47. GFFT: Global-local feature fusion transformers for facial expression

摘要: Facial expression recognition in the wild has become more challenging owing to various unconstrained conditions, such as facial occlusion and pose variation. Previous methods usually recognize expressions by holistic or relatively coarse local methods, but only capture limited features and are susceptible to be influenced. In this paper, we propose the Global-local Feature Fusion Transformers (GFFT) that is centered on cross-patch communication between features by self-attentive fusion. This method solves the problems of facial occlusion and pose variation effectively. Firstly, the Global Contextual Information Perception (GCIP) is designed to fuse global and local features, learning the relationship between them. Subsequently, the Facial Salient Feature Perception (FSFP) module is proposed to guide the fusion features to understand the key regions of facial features using facial landmark features to further capture face-related salient features. In addition, the Multi-scale Feature Fusion (MFF) is constructed to combine different stages of fusion features to reduce the sensitivity of the deep network to facial occlusion. Extensive experiments show that our GFFT outperforms existing state-of-the-art methods with 92.05% on RAF-DB, 67.46% on AffectNet-7, 63.62% on AffectNet-8, and 91.04% on FERPlus, demonstrating its effectiveness and robustness.  
Web of Science 核心合集中的 "被引频次":  
8  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
0262-8856  
eISSN:  
1872-8138  
入藏号:  
WOS:001088742600001  
Indexed Date:  
2023-11-08  
Institution Address:  
Hangzhou Dianzi Univ, Sch Media & Design, Hangzhou 310018, Peoples R China  
Institution Address:  
Hangzhou Dianzi Univ, Wenzhou Inst, Wenzhou 325038, Peoples R China  
Institution Address:  
China Univ Geosci, Sch Comp Sci, Wuhan 430074, Hubei, Peoples R China  
Institution:  
Institution:  
Institution:

### 48. Efficient Metal Corrosion Area Detection Model Combining Convolution and

摘要: In the context of rapid industrialization, efficiently detecting metal corrosion areas has become a critical task in preventing material damage. Unlike conventional semantic segmentation targets, metal corrosion characteristics vary significantly in color, texture, and size. Traditional image segmentation methods need improvement in scenarios involving occlusions, shadows, and defects. This paper proposes a convolution and sequence encoding combined network, MCD-Net, for metal corrosion area segmentation. First, a visual Transformer sequence encoder is introduced into the convolutional encoder-decoder network to enhance global information processing capabilities and establish long-range feature dependencies. A feature fusion method based on an attention module is proposed to enhance the model's ability to recognize corrosion boundaries, thereby enhancing segmentation accuracy and model robustness. Finally, in the model's decoding stage, a score-based multi-scale feature enhancement method is employed to emphasize significant features in the corrosion areas. Experimental results indicate that this method attained an F1 score of 84.53% on a public corrosion dataset, demonstrating the model's deeper understanding and reasoning capabilities for shadow and defect features, as well as excellent noise resistance performance.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001351052500001  
Indexed Date:  
2024-11-16  
Institution Address:  
Nantong Univ, Sch Elect Engn & Automat, Nantong 226001, Peoples R China  
Institution:

### 49. Self-attention Guidance Based Crowd Localization and Counting

摘要: Most existing studies on crowd analysis are limited to the level of counting, which cannot provide the exact location of individuals. This paper proposes a self-attention guidance based crowd localization and counting network (SA-CLCN), which can simultaneously locate and count crowds. We take the form of object detection, using the original point annotations of crowd datasets as supervision to train the network. Ultimately, the center point coordinate of each head as well as the number of crowds are predicted. Specifically, to cope with the spatial and positional variations of the crowd, the proposed method introduces transformer to construct a globallocal feature extractor (GLFE) together with the convolutional structure. It establishes the near-to-far dependency between elements so that the global context and local detail features of the crowd image can be extracted simultaneously. Then, this paper designs a pyramid feature fusion module (PFFM) to fuse the global and local information from high level to low level to obtain a multiscale feature representation. In downstream tasks, this paper predicts candidate point offsets and confidence scores by a simple regression header and classification header. In addition, the Hungarian algorithm is used to match the predicted point set and the labelled point set to facilitate the calculation of losses. The proposed network avoids the errors or higher costs associated with using traditional density maps or bounding box annotations. Importantly, we have conducted extensive experiments on several crowd datasets, and the proposed method has produced competitive results in both counting and localization.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
被引频次合计:  
1  
ISSN:  
2731-538X  
eISSN:  
2731-5398  
入藏号:  
WOS:001169576500005  
Indexed Date:  
2024-03-19  
Institution Address:  
Yanshan Univ, Sch Informat Sci & Engn, Qinhuangdao 066000, Peoples R China  
Institution Address:  
Hebei Key Lab Informat Transmiss & Signal Proc, Qinhuangdao 066000, Peoples R China  
Institution:  
Institution:  
Hebei Key Lab Informat Transmiss & Signal Proc

### 50. Lightweight multi-target detection algorithm for unmanned aerial vehicle

摘要: Compared with the image captured in the natural scene, the image obtained by unmanned aerial vehicle (UAV) aerial photography has a more complex background and many dense small targets, which puts forward higher requirements for the detection accuracy of the target detection algorithm. However, because the UAV is a kind of small mobile device, how to ensure its real-time detection effect has been a problem. Aiming at these problems, the lightweight YOLOv7 algorithm, namely LRT-YOLOv7, is designed. First, the enhance feature fusion module and the transformer efficient layer aggregation networks module are proposed to improve the performance of feature extraction and fusion to enhance the efficiency of small target detection. Second, aiming at the problems of small target size and complex background in the UAV images, the detection head structure is redesigned in the YOLOv7-tiny algorithm to enhance the multi-scale feature fusion ability of the algorithm and thereby improve the algorithm's detection accuracy for small targets. Finally, ablation, comparison, and visualization validation experiments were conducted using precision, recall, mean average precision, and frames per second (FPS) as evaluation indicators. The results show that the detection speed of the LRT-YOLOv7 algorithm on the self-made traffic target dataset is 133.8 FPS, and the precision indicator is 84.58%. Therefore, the LRT-YOLOv7 algorithm has high accuracy and real-time performance in traffic target detection tasks for UAV aerial imagery.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
1931-3195  
入藏号:  
WOS:001134904400008  
Indexed Date:  
2024-02-25  
Institution Address:  
Northeastern Univ, Sch Mech Engn & Automat, Shenyang, Peoples R China  
Institution Address:  
Shenyang Inst Engn, Coll Energy & Power, Shenyang, Peoples R China  
Institution:  
Institution:

### 51. Multi-Scale Feature Fusion Based on PVTv2 for Deep Hash Remote Sensing

摘要: For high-resolution remote sensing image retrieval tasks, single-scale features cannot fully express the complexity of the image information. Due to the large volume of remote sensing images, retrieval requires extensive memory and time. Hence, the problem of how to organically fuse multi-scale features and enhance retrieval efficiency is yet to be resolved. We propose an end-to-end deep hash remote sensing image retrieval model (PVTA\_MSF) by fusing multi-scale features based on the Pyramid Vision Transformer network (PVTv2). We construct the multi-scale feature fusion module (MSF) by using a global attention mechanism and a multi-head self-attention mechanism to reduce background interference and enhance the representation capability of image features. Deformable convolution is introduced to address the challenge posed by varying target orientations. Moreover, an intra-class similarity (ICS) loss is proposed to enhance the discriminative capability of the hash feature by minimizing the distance among images of the same category. The experimental results show that, compared with other state-of-the-art methods, the proposed hash feature could yield an excellent representation of remote sensing images and improve remote sensing image retrieval accuracy. The proposed hash feature can gain an increase of 4.2% and 1.6% in terms of mAP on the UC Merced and NWPU-RESISC45 datasets, respectively, in comparison with other methods.  
Web of Science 核心合集中的 "被引频次":  
5  
在中国科学引文数据库中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
eISSN:  
2072-4292  
入藏号:  
WOS:001081181200001  
Indexed Date:  
2023-10-28  
Institution Address:  
East China Univ Technol, Key Lab Mine Environm Monitoring & Improving Poya, Minist Nat Resources, Nanchang 330013, Peoples R China  
Institution Address:  
East China Univ Technol, Sch Surveying & Geoinformat Engn, Nanchang 330013, Peoples R China  
Institution Address:  
Geog Informat Engn Brigade Geol Bur Jiangxi Prov, Nanchang 330002, Peoples R China  
Institution:  
Institution:  
Institution:  
Geog Informat Engn Brigade Geol Bur Jiangxi Prov

### 52. Waste classification strategy based on multi-scale feature fusion for

摘要: Waste classification is an important measure to protect the environment. Existing waste classification methods mainly focus on scientific research, but lack attention to the challenges of waste classification in actual scenarios. For example, wastes with similar contours, similar textures, or contaminated appearance are difficult to be classified in actual scenarios. To address these issues, this paper proposes an innovative multi-scale feature fusion strategy (MFFS) to improve the classification accuracy of these wastes. MFFS combines local fine-grained features with global coarse-grained features to improve the feature expression ability of waste. However, how to effectively fuse these two features is a key challenge. This paper proposes a dual-scale feature fusion strategy, first fusing fine-grained features in the first dimension, then fusing coarse-grained features in the second dimension, and introducing spatial features to further enhance feature expression capabilities. In order to reduce the interference of background information, the model in this paper models global relationships based on convolutional features. The MFFS strategy achieved a classification accuracy of 95.5% on the self-built dataset and 94.1% on the public dataset TrashNet. The number of parameters of our model is reduced by 57.2% compared with the classic VGG16 and by 34.2% compared with the Vision Transformer. In addition, we designed an intelligent waste sorting device and deployed the MFFS model on the device to implement the application. Experiments show that our model has ideal accuracy and stability and can be promoted and applied.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0956-053X  
eISSN:  
1879-2456  
入藏号:  
WOS:001339001400001  
PubMed ID:  
39418863  
Indexed Date:  
2024-10-27  
Institution Address:  
Hangzhou Dianzi Univ, Sch Mech Engn, Hangzhou 310018, Peoples R China  
Institution Address:  
Zhejiang Univ, State Key Lab Fluid Power Mech Syst, Hangzhou 310027, Peoples R China  
Institution Address:  
Hangzhou City Univ, Dept Mech Engn, Hangzhou 310015, Peoples R China  
Institution:  
Institution:  
Institution:

### 53. EF-DETR: A Lightweight Transformer-Based Object Detector With an

摘要: Object detection plays a key role in helping to enable industrial quality control and safety monitoring. This article introduces a lightweight and efficient transformer-based object detection network called the encoder-free DEtection TRansformer (EF-DETR). This novel architecture enhances the DETR model through a redesigned network structure, leading to improved accuracy in object detection and a more lightweight network. To address the issue of suboptimal object detection accuracy, especially for small objects in the DETR model, we introduce a multiscale feature extractor and a high-efficiency feature fusion module. These components facilitate the direct extraction of fine-grained features, thereby enabling effective object detection. Departing from the use of a high-complexity encoder structure, we explore the utilization of an encoder-free neck structure to reduce the network's computational complexity. In addition, to expedite convergence, denoising training is incorporated into the decoder. This article presents extensive experiments, and the EF-DETR demonstrates strong performance on the MS COCO2017 dataset compared to other popular models.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1551-3203  
eISSN:  
1941-0050  
入藏号:  
WOS:001279054900001  
Indexed Date:  
2024-08-04  
Institution Address:  
Chongqing Univ, Sch Comp Sci, Chongqing 400044, Peoples R China  
Institution Address:  
Chongqing Univ, State Key Lab Mech Transmiss, Chongqing 400044, Peoples R China  
Institution Address:  
Beijing Normal Univ Zhuhai, BNU UIC Inst Artificial Intelligence & Future Netw, Zhuhai 519087, Peoples R China  
Institution Address:  
BNU HKBU United Int Coll Zhuhai, Guangdong Key Lab AI Multimodal Data Proc, Zhuhai 519087, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
BNU HKBU United Int Coll Zhuhai

### 54. Transformer-based cross-modality interaction guidance network for RGB-T

摘要: Exploring more effective multimodal fusion strategies is still challenging for RGB-T salient object detection (SOD). Most RGB-T SOD methods tend to focus on the strategy of acquiring modal complementary features by utilizing foreground information while ignoring the importance of background information for salient object localization. In addition, feature fusion without information filtering may introduce more noise. To solve these problems, this paper proposes a new cross-modal interaction guidance network (CIGNet) for RGB-T saliency object detection. Specifically, we construct a transformer-based dual-stream encoder to extract multimodal features. In the decoder, we propose an attention mechanism-based modal information complementary module (MICM) for capturing cross-modal complementary information for global comparison and salient object localization. Based on the MICM features, we design a multi-scale adaptive fusion module (MAFM) to find the optimal salient region of the multi-scale fusion process and reduce redundant features. In order to enhance the completeness of salient features after multi-scale feature fusion, this paper proposes the saliency region mining module (SRMM), which corrects the features in the boundary neighborhood by exploiting the differences between foreground and background pixels and the boundary. Comparisons with other state-of-the-art methods on three RGB-T datasets and five RGB-D datasets, the experimental results demonstrate the superiority and extensiveness of the proposed CIGNet.  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:001269381600001  
Indexed Date:  
2024-07-22  
Institution Address:  
Henan Univ, Sch Phys & Elect, Kaifeng 475004, Peoples R China  
Institution:

### 55. Enhancing target detection accuracy through cross-modal spatial

摘要: The disparity between human and machine perception of spatial information presents a challenge for machines to accurately sense their surroundings and improve target detection performance. Cross-modal data fusion emerges as a potential solution to enhance the perceptual capabilities of systems. This article introduces a novel spatial perception method that integrates dual-modality feature fusion and coupled attention mechanisms to validate the improvement in detection performance through cross-modal information fusion. The proposed approach incorporates cross-modal feature extraction through a multi-scale feature extraction structure employing a dual-flow architecture. Additionally, a transformer is integrated for feature fusion, while the information perception of the detection system is optimized through the utilization of a linear combination of loss functions. Experimental results demonstrate the superiority of our algorithm over single-modality target detection using visible images, exhibiting an average accuracy improvement of 30.4%. Furthermore, our algorithm outperforms single-modality infrared image detection by 3.0% and comparative multimodal target detection algorithms by 3.5%. These results validate the effectiveness of our proposed algorithm in fusing dual-band features, significantly enhancing target detection accuracy. The adaptability and robustness of our approach are showcased through these results.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2296-424X  
入藏号:  
WOS:001250325300001  
Indexed Date:  
2024-06-25  
Institution Address:  
Chinese Acad Sci, Shanghai Inst Tech Phys, Shanghai, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing, Peoples R China  
Institution Address:  
Chinese Acad Sci, Key Lab Infrared Syst Detect & Imaging Technol, Shanghai, Peoples R China  
Institution:  
Institution:  
Institution:

### 56. TAMDepth: self-supervised monocular depth estimation with transformer

摘要: Self-supervised monocular depth estimation presents a promising result, which utilizes image sequences instead of challenging-to-source ground truth for training. The framework of most current studies on self-supervised depth estimation is based on fully convolutional or transformer architectures, and there is little discussion on the hybrid architecture. In this paper, we proposed TAMDepth, a novel framework that can effectively capture the local and global features of image sequences by combining convolutional blocks and transformer blocks. TAMDepth adopts multi-scale feature fusion convolutional modules capture local details in shallow layers while transformer blocks build the global dependency in higher layers. Furthermore, to enhance the representation of architecture, we introduce an adapter modulation that injects the spatial prior to the transformer blocks through cross-attention, which improves the ability of modeling the scene. Experiments demonstrate that our model exhibits state-of-the-art performance on the KITTI dataset and also shows strong generalization performance on the Make3D dataset. Source code is available at https://github.com/deansaice/TAMDepth.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0178-2789  
eISSN:  
1432-2315  
入藏号:  
WOS:001194660900001  
Indexed Date:  
2024-04-07  
Institution Address:  
Hubei Univ Automot Technol, Inst Automot Engineers, 167 Checheng West Rd, Shiyan 442000, Hubei, Peoples R China  
Institution:

### 57. Multiscale Hyperspectral Pansharpening Network Based on Dual Pyramid and

摘要: Hyperspectral pansharpening is to fuse a high spatial resolution panchromatic image (PAN) with a low spatial resolution hyperspectral image (LR-HSI) and generate high resolution hyperspectral image (HR-HSI). However, most existing deep learning-based pansharpening methods have some issues, such as spectral distortion and insufficient spatial texture enhancement. In this work, we propose a novel multiscale pansharpening network based on the Dual Gaussian-Laplacian Pyramid (DGLP) and Transformer, named MDTP-Net. Specifically, the DGLP module is designed to obtain feature maps at multilevel scales, which effectively learn global spectral information and spatial detail texture information. Then, we design a corresponding Transformer module for each scale feature and utilize the multihead attention mechanism to guide the extraction of spatial information from LR-HSI and PAN images. This enhances the stability of pansharpening and improves the fusion of spectral with spatial information across feature spaces. In addition, the feature extractors are inserted to connect DGLP and Transformer, making the spatial feature map smoother and richer in channel and texture features. The feature fusion and multiscale feature connection blocks are used to connect multiscale information together to generate HR-HSI images with more comprehensive spatial and spectral features. Finally, extensive experiments on three classic hyperspectral datasets are conducted. The experimental results demonstrate that our proposed MDTP-Net outperforms conventional methods and existing deep learning-based methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
被引频次合计:  
2  
ISSN:  
1939-1404  
eISSN:  
2151-1535  
入藏号:  
WOS:001355784300015  
Indexed Date:  
2024-11-21  
Institution Address:  
Beijing Univ Civil Engn & Architecture, Sch Sci, Beijing 100044, Peoples R China  
Institution Address:  
Beijing Univ Civil Engn & Architecture, Inst Big Data Modeling & Technol, Beijing 100044, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution:  
Institution:  
Institution:

### 58. SwinPA-Net: Swin Transformer-Based Multiscale Feature Pyramid

摘要: The precise segmentation of medical images is one of the key challenges in pathology research and clinical practice. However, many medical image segmentation tasks have problems such as large differences between different types of lesions and similar shapes as well as colors between lesions and surrounding tissues, which seriously affects the improvement of segmentation accuracy. In this article, a novel method called Swin Pyramid Aggregation network (SwinPA-Net) is proposed by combining two designed modules with Swin Transformer to learn more powerful and robust features. The two modules, named dense multiplicative connection (DMC) module and local pyramid attention (LPA) module, are proposed to aggregate the multiscale context information of medical images. The DMC module cascades the multiscale semantic feature information through dense multiplicative feature fusion, which minimizes the interference of shallow background noise to improve the feature expression and solves the problem of excessive variation in lesion size and type. Moreover, the LPA module guides the network to focus on the region of interest by merging the global attention and the local attention, which helps to solve similar problems. The proposed network is evaluated on two public benchmark datasets for polyp segmentation task and skin lesion segmentation task as well as a clinical private dataset for laparoscopic image segmentation task. Compared with existing state-of-the-art (SOTA) methods, the SwinPA-Net achieves the most advanced performance and can outperform the second-best method on the mean Dice score by 1.68%, 0.8%, and 1.2% on the three tasks, respectively.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
9  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
46  
被引频次合计:  
46  
eISSN:  
2162-2388  
入藏号:  
MEDLINE:36121961  
PubMed ID:  
36121961  
Indexed Date:  
2024-06-19

### 59. MSMT-LCL: Multiscale Spatial-Spectral Masked Transformer With Local

摘要: Deep learning plays a crucial role in hyperspectral image (HSI) classification, with the Transformer being highly favored by researchers due to its exceptional ability to model long-range dependencies. However, the Transformer necessitates a substantial amount of labeled training samples to train its numerous parameters, exacerbating the challenge of training an effective HSI classification Transformer model, particularly given the inherent scarcity of HSI data. Therefore, we propose a novel method for HSI classification, termed multiscale spatial-spectral masked Transformer with local contrastive learning (MSMT-LCL). This method consists of two stages: self-supervised pretraining and supervised fine-tuning. Initially, we utilize the multiscale augmented feature mapping module (MAFM) to project original HSI data into two mixed-scale feature maps, which are then separately fed into two masked Transformer branches for reconstruction. To facilitate the model in learning the dependency relationships between central pixel land-cover information and neighboring land cover, we introduce a novel mask strategy based on center-patch. Furthermore, in the pretraining stage, we integrate local contrastive learning (LCL) to enable the model to focus on local center information at varying scales. Upon completion of pretraining, the network undergoes fine-tuning to obtain feature maps at two different scales. Subsequently, we devise a novel adaptive multiscale feature fusion module (AMFM) to adaptively aggregate these two features and produce the final classification results. Extensive experiments on three real datasets demonstrate the superiority of our proposed MSMT-LCL method over several state-of-the-art HSI classification methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001338406700037  
Indexed Date:  
2024-11-23  
Institution Address:  
East China Jiaotong Univ, Sch Informat & Software Engn, Nanchang 330013, Peoples R China  
Institution Address:  
Guangzhou Univ, Sch Elect & Commun Engn, Guangzhou 511370, Peoples R China  
Institution Address:  
Hubei Univ, Hubei Key Lab Appl Math, Key Lab Intelligent Sensing Syst & Secur, Minist Educ, Wuhan 430062, Peoples R China  
Institution Address:  
Hubei Univ, Fac Math & Stat, Wuhan 430062, Peoples R China  
Institution Address:  
KTH Royal Inst Technol, Sch Architecture & Built Environm, Div Geoinformat, S-11428 Stockholm, Sweden  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 60. UGTransformer: A Sheep Extraction Model From Remote Sensing Images for

摘要: The extraction of sheep from satellite images plays an extremely important role in the precise automation of animal husbandry management. Current methods of extracting sheep mainly use hardware, such as radio frequency equipment and visual ear tags, which are prone to loss or damage. In this study, a new network, UGTransformer, was developed to extract sheep from high spatial resolution remote sensing (RS) images. In UGTransformer, a merge block was designed to fuse two scales of features in the encoder to improve the multiscale feature fusion capability. It enhanced the integration of global context features and spatial detailed features by combining the features in the decoder. A global connectivity module containing two sliding sub-modules, horizontal and vertical, was developed to correlate the horizontal and vertical features and correlate the arbitrary positions of the feature maps through the integration of the two modules, which realized the extraction of global contextual information. Our experimental results showed that the proposed UGTransformer performed well in comparison with UNet, Deeplab v3+, DCSwin, BANet, and UNetFormer, four recently proposed network structures for semantic segmentation. UGTransformer achieved at least a 1.8% increase in mean intersection over the union. This study not only provided potential solutions for the problems inherent in large-scale sheep extraction but also developed mechanisms for small-object extraction. The implementation code is available at https://github.com/chenchengStore/GlobalLocalAttention, and the RS images used in this study are available at https://github.com/chencheng-2023/UGTransformer-remote-sensing-images.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
被引频次合计:  
6  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001164472500040  
Indexed Date:  
2024-03-21  
Institution Address:  
Guilin Univ Elect Technol, Sch Comp Sci & Informat Secur, Guilin 541004, Peoples R China  
Institution Address:  
Int Res Ctr Big Data Sustainable Dev Goals, Beijing 100094, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Key Lab Digital Earth Sci, Beijing 100094, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Coll Resources & Environm, Beijing 100049, Peoples R China  
Institution Address:  
Peking Univ, Coll Urban & Environm Sci, Beijing 100871, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Geog Sci & Nat Resources Res, State Key Lab Resources & Environm Informat Syst, Beijing 100101, Peoples R China  
Institution Address:  
Minist Ecol & Environm, Satellite Applicat Ctr Ecol & Environm, Beijing 100094, Peoples R China  
Institution Address:  
State Environm Protect Key Lab Satellite Remote Se, Beijing 100094, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Minist Ecol & Environm  
Institution:  
State Environm Protect Key Lab Satellite Remote Se

### 61. PSO-YOLO: a contextual feature enhancement method for small object

摘要: In recent years, UAV aerial object detection has been widely applied in fields such as urban planning, environmental monitoring, agricultural management, and disaster assessment. However, challenges such as low accuracy, missed detections, and occlusion still persist when detecting small objects from aerial perspectives. This study proposes the PSO-YOLO (Precise Small Object-You Only Look Once) algorithm, which is developed based on multi-scale feature fusion and extraction techniques for small object detection. Firstly, the MSFE (Multi-Scale Feature Extraction) module was developed, leveraging Transformer-based multi-scale feature fusion and self-attention mechanisms to enhance the acquisition of fine-grained multi-scale details. Secondly, a small object-enhanced STE-Neck (Small Object Enhancement-Neck) network is introduced to fuse multi-scale features during downsampling, effectively capturing detailed and local features to prevent missed detections and feature loss due to occlusion. The fused features are then input into a small object detection head. Finally, the Swin Transformer is incorporated into the YOLOv8n backbone network to improve the C2f module, thereby enhancing its ability to capture long-range contextual information and further improve small object feature extraction. Additionally, the SPPF-LSKA (Spatial Pyramid Pooling Fast-Large Separable Kernel Attention) module is introduced to replace the SPPF (Spatial Pyramid Pooling Fast) module, aiming to enhance the algorithm's ability to extract key features. Experiments on the VisDrone2019 dataset show that the PSO-YOLO algorithm achieves an mAP@50% of 30.9% and an mAP@50-95% of 17.4%, representing improvements of 4.6% and 2.8%, respectively, over YOLOv8n.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1865-0473  
eISSN:  
1865-0481  
入藏号:  
WOS:001421393600003  
Indexed Date:  
2025-02-20  
Institution Address:  
Shijiazhuang Tiedao Univ, Sch Informat Sci & Technol, Shijiazhuang 050043, Peoples R China  
Institution Address:  
Shijiazhuang Tiedao Univ, State Key Lab Mech Behav & Syst Safety Traff Engn, Shijiazhuang 050043, Peoples R China  
Institution:  
Institution:

### 62. Human action recognition method based on multi-scale feature fusion of

摘要: In order to solve the problem of insufficient mining of potential association between remote nodes in human action recognition tasks, and the problem of high training cost caused by using multi-modal data, a multi-scale feature fusion human action recognition method under the condition of single mode was proposed. Firstly, the global feature correlation of the original skeleton diagram of human body was carried out, and the coarse-scale global features were used to capture the connections between the remote nodes. Secondly, the global feature correlation graph was divided locally to obtain the Complementary Subgraphs with Global Features (CSGFs), the fine-scale features were used to establish the strong correlation, and the multi-scale feature complementarity was formed. Finally, the CSGFs were input into the spatialtemporal Graph Convolutional module for feature extraction, and the extracted results were aggregated to output the final classification results. Experimental results show that the accuracy of the proposed method on the authoritative action recognition dataset NTU RGB+D60 is 89.0% (X-sub) and 94.2% (X-view) respectively. On the challenging large-scale dataset NTU RGB+D120, the accuracy of the proposed method is 83.3% (X-sub) and 85.0% (X-setup) respectively, which is 1.4 and 0.9 percentage points higher than that of the ST-TR (Spatial-Temporal TRansformer) under single modal respectively, and 4.1 and 3.5 percentage points higher than that of the lightweight SGN (Semantics-Guided Network). It can be seen that the proposed method can fully exploit the synergistic complementarity of multi-scale features, and effectively improve the recognition accuracy and training efficiency of the model under the condition of single modal.  
摘要:  
针对人体行为识别任务中未能充分挖掘超距关节点之间潜在关联的问题,以及使用多模态数据带来的高昂训练成本的问题,提出一种单模态条件下的多尺度特征融合人体行为识别方法。首先,将人体的原始骨架图进行全局特征关联,并利用粗尺度的全局特征捕获远距离关节点间的联系;其次,对全局特征关联图进行局部划分以得到融合了全局特征的互补子图(CSGF),利用细尺度特征建立强关联,并形成多尺度特征的互补;最后,将CSGF输入时空图卷积模块中提取特征,并聚合提取后的结果以输出最终的分类结果。实验结果表明,在行为识别权威数据集NTU RGB+D60上,所提方法的准确率分别为89.0%(X-sub)和94.2%(X-view);在具有挑战性的大规模数据集NTU RGB+D120上,所提方法的准确率分别为83.3%(X-sub)和85.0%(X-setup),与单模态下的ST-TR(Spatial-Temporal TRansformer)相比,分别提升1.4和0.9个百分点,与轻量级SGN(Semantics-Guided Network)相比,分别提升4.1和3.5个百分点。可见,所提方法能够充分挖掘多尺度特征的协同互补性,并有效提高单模态条件下模型的识别准确率和训练效率。  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1001-9081  
入藏号:  
CSCD:7579969  
Indexed Date:  
2024-01-06  
Institution Address:  
常州大学计算机与人工智能学院阿里云大数据学院软件学院;;南京理工大学, ;;江苏省社会安全图像与视频理解重点实验室, 常州;;南京, 江苏;; 213164;;210094, 中国  
Institution Address:  
常州大学计算机与人工智能学院阿里云大数据学院软件学院, 常州, 江苏 213164, 中国  
Institution Address:  
School of Computer Science and Artificial Intelligence,Aliyun School of Big Data,School of Software,Changzhou University;;Nanjing University of Science and Technology, ;;Jiangsu Key Laboratory of Image and Video Understanding for Social Security, Changzhou;;Nanjing, Jiangsu;;Jiangsu 213164;;210094  
Institution Address:  
School of Computer Science and Artificial Intelligence,Aliyun School of Big Data,School of Software,Changzhou University, Changzhou, Jiangsu 213164, China  
Institution:  
常州大学计算机与人工智能学院阿里云大数据学院软件学院;;南京理工大学  
Institution:  
常州大学计算机与人工智能学院阿里云大数据学院软件学院  
Institution:  
School of Computer Science and Artificial Intelligence,Aliyun School of Big Data,School of Software,Changzhou University;;Nanjing University of Science and Technology  
Institution:  
School of Computer Science and Artificial Intelligence,Aliyun School of Big Data,School of Software,Changzhou University

### 63. Infrared Image Detection of Substation Insulators Using an Improved

摘要: The thermographic diagnostic of substation insulators can facility to detect their insulation status in time. However, the detection effect of infrared insulator images is usually poor due to the complex background of substation scene images. In order to solve this problem and improve the automatic detection capability of substation equipment, an infrared insulator image detection model based on the improved feature fusion single shot multibox detector is proposed in this paper. The model combines multi-scale feature maps to generate a new feature pyramid, and designs a new feature enhancement module in the shallow network of the model to improve its ability to extract features of infrared images of substation insulators. Meanwhile, the clustering algorithm is used to calculate the target aspect ratio information of the dataset to realize the adaptive change of the aspect ratio of the default box. Furthermore, the transfer learning is employed to further improve the learning efficiency of the model. Before the formal experiment, the design scheme of feature enhancement module is confirmed by ablation studies. Experiments are performed on the collected infrared insulator images, instrument transformer images and Pascal VOC 2007 dataset. It is demonstrated experimentally that this model achieves excellent detection results.  
在中国科学引文数据库中的被引频次:  
6  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
46  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
52  
ISSN:  
0885-8977  
eISSN:  
1937-4208  
入藏号:  
WOS:000722005300011  
Indexed Date:  
2021-12-01  
Institution Address:  
Guangxi Univ, Sch Elect Engn, Nanning 530004, Peoples R China  
Institution Address:  
State Grid Zibo Elect Power Supply Co, Zibo 255000, Peoples R China  
Institution:  
Institution:  
State Grid Zibo Elect Power Supply Co

### 64. A multi-scale semantic feature fusion method for remote sensing crop

摘要: The efficient and accurate acquisition of crop spatial planting structure information is of great significance to ensure national food security. Current crop classification studies have still not achieved optimal results in accurate crop classification identification due to the complex heterogeneous details of clear images and insufficient analysis and utilization of information. This study takes the Yuncheng-Linfen Basin in Shanxi, China as the study area and uses the ReliefF-RFE feature selection algorithm to achieve data dimensionality reduction. Contextual semantic feature aggregation, spatial channel attention mechanism, and hierarchical refinement strategy are used to fuse the multi-scale feature information in the extracted remote sensing images from the top down and bottom up. This study shows that 1) the ReliefF-RFE algorithm can effectively reduce the dimensionality of 224dimensional multisource features and generate a 31-dimensional optimal feature space subset with strong interpretation to improve the computational efficiency of the model; 2) the multi-scale feature fusion model constructed in this study is better than UNet, ResNet, DeepLabv3+, HRNet and Swin Transformer in classification recognition of winter wheat and corn, with an overall accuracy of 95.76% and 93.44%; 3) A multiple feature fusion strategy designed through ablation experiments was evaluated to evaluate the reasonableness of each module and validate the irreplaceability and superiority of the model for improving crop accuracy in complex and heterogeneous terrain; 4) the multi-scale feature fusion model converts deep abstract features into local detail information by cross-scale contextual aggregation of deep and shallow multi-scale features, it obtains good extraction effect on fragmented small plots, and the reduction of local fine details of features is significantly better than other models. The multi-scale feature fusion classification model constructed in this study fully considers the interpretability of the input features and fuses the semantic features at different scales to eliminate the "semantic gap", which shows that the model has strong stability and robustness and provides a reference for fast and accurate mapping of Sentinel-2 images of crops in large areas.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
3  
ISSN:  
0168-1699  
eISSN:  
1872-7107  
入藏号:  
WOS:001260000700001  
Indexed Date:  
2024-07-08  
Institution Address:  
Beijing Normal Univ, Fac Geog Sci, Beijing 100875, Peoples R China  
Institution:

### 65. Research on Intelligent Detection Algorithms for Substation Power

摘要: It is difficult for mobile edge devices to quickly detect high-similarity targets in complex environments during substation power meter inspections. To address this, we proposed a lightweight YOLOX-based method for power meter image detection. First, we built the YOLOX detection network and designed a depthwise separable convolutional backbone feature extraction structure and a multi-scale feature fusion structure based on parameter reorganization to compress model parameters and improve inference speed. Secondly, a three-dimensional attention mechanism, SimAM, was embedded in the feature fusion layer. This mechanism learns the energy distribution of features and weights the target areas to enhance meter detection in complex environments. Additionally, to address specific issues in power meter detection, we designed a transformer structure based on pyramid pooling feature encoding, focusing on refining local features and capturing long-distance features to mine high-semantic information, thereby improving the detection accuracy of power meters with different shapes. Finally, we validated the algorithm by constructing a dataset of broken, blurred, and normal power pointer meters. The experimental results show that the improved model increases the mean precision from 75.49% to 85.93% and the detection speed from 36 Frame/s to 45 Frame/s compared to the original model. On the mobile hardware Jetson NX, the inference speed reaches 17.6 Frame/s. Compared to other lightweight models, this model has significant advantages in detection accuracy and speed, providing a feasible technical solution for the visualization, informatization, and intelligence of power meters.  
摘要:  
移动边缘端设备在变电站电力仪表检测中,难以快速检测复杂环境中高相似度目标。为此,提出了一种基于轻量级YOLOX网络的电力仪表图像检测方法。首先,搭建YOLOX检测网络,并设计了基于深度可分离卷积骨干特征提取结构和参数重组的多尺度特征融合结构,以压缩模型参数量和提升推理速度。其次,在特征融合层中嵌入3维注意力机制SimAM,通过学习特征的能量分布,对目标区域进行加权,提升复杂环境下的仪表检测能力。同时,针对电力仪表检测的特殊需求,设计了基于金字塔池化特征编码的Transformer结构,从局部特征细化和长距离特征捕获2个方面挖掘底层高语义信息特征,提高不同外形电力仪表的检测精度。最后,通过构建破损、模糊及正常3种类型的电力指针型仪表数据集进行验证。实验结果显示,改进的模型相比原始模型,均值精度从75.49%提升至85.93%,检测速度从36帧/s提升至45帧/s。在移动端硬件Jetson NX上,推理速度可达17.6帧/s。与其他轻量化模型相比,该模型在检测精度和速度上具有明显优势,为电力仪表的可视化、信息化和智能化提供了可行的技术方案。  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1003-6520  
入藏号:  
CSCD:7769332  
Indexed Date:  
2024-09-27  
Institution Address:  
综合能源电力装备及系统安全湖北省重点实验室;;武汉大学电气与自动化学院, 综合能源电力装备及系统安全湖北省重点实验室;;, 武汉;;武汉, ;; 430072;;430072  
Institution Address:  
国网信息通信产业集团有限公司, 北京 102211, 中国  
Institution Address:  
海军工程大学, 舰船综合电力技术国防科技重点实验室, 武汉, 湖北 430034, 中国  
Institution Address:  
Hubei Key Laboratory of Power Equipment & System Security for Integrated Energy;;School of Electrical and Automation, Wuhan University, Hubei Key Laboratory of Power Equipment & System Security for Integrated Energy;;, Wuhan;;Wuhan, ;; 430072;;430072  
Institution Address:  
State Grid Information & Telecommunication Group Co., Ltd., Beijing 102211, China  
Institution Address:  
Naval University of Engineering, National Key Laboratory of Science & Technology on Vessel Integrated Power System, Wuhan, Hubei 430034, China  
Institution:  
综合能源电力装备及系统安全湖北省重点实验室;;武汉大学电气与自动化学院  
Institution:  
国网信息通信产业集团有限公司  
Institution:  
海军工程大学  
Institution:  
Hubei Key Laboratory of Power Equipment & System Security for Integrated Energy;;School of Electrical and Automation, Wuhan University  
Institution:  
State Grid Information & Telecommunication Group Co., Ltd.  
Institution:  
Naval University of Engineering

### 66. Detection Transformer with Multi-Scale Fusion Attention Mechanism for

摘要: Casting defects in turbine blades can significantly reduce an aero-engine's service life and cause secondary damage to the blades when exposed to harsh environments. Therefore, casting defect detection plays a crucial role in enhancing aircraft performance. Existing defect detection methods face challenges in effectively detecting multi-scale defects and handling imbalanced datasets, leading to unsatisfactory defect detection results. In this work, a novel blade defect detection method is proposed. This method is based on a detection transformer with a multi-scale fusion attention mechanism, considering comprehensive features. Firstly, a novel joint data augmentation (JDA) method is constructed to alleviate the imbalanced dataset issue by effectively increasing the number of sample data. Then, an attention-based channel-adaptive weighting (ACAW) feature enhancement module is established to fully apply complementary information among different feature channels, and further refine feature representations. Consequently, a multi-scale feature fusion (MFF) module is proposed to integrate high-dimensional semantic information and low-level representation features, enhancing multi-scale defect detection precision. Moreover, R-Focal loss is developed in an MFF attention-based DEtection TRansformer (DETR) to further solve the issue of imbalanced datasets and accelerate model convergence using the random hyper-parameters search strategy. An aero-engine turbine blade defect X-ray (ATBDX) image dataset is applied to validate the proposed method. The comparative results demonstrate that this proposed method can effectively integrate multi-scale image features and enhance multi-scale defect detection precision.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
eISSN:  
1424-8220  
入藏号:  
WOS:001183111400001  
PubMed ID:  
38475200  
Indexed Date:  
2024-03-19  
Institution Address:  
Jiangsu Univ Sci & Technol, Sch Mech Engn, Zhenjiang 212003, Peoples R China  
Institution:

### 67. MC-Net: multi-scale contextual information aggregation network for image

摘要: Remote Sensing Image Captioning (RSIC) plays a crucial role in advancing semantic understanding and has increasingly become a focal point of research. Nevertheless, existing RSIC methods grapple with challenges due to the intricate multi-scale nature and multifaceted backgrounds inherent in Remote Sensing Images (RSIs). Compounding these challenges are the perceptible information disparities across diverse modalities. In response to these challenges, we propose a novel multi-scale contextual information aggregation image captioning network (MC-Net). This network incorporates an image encoder enhanced with a multi-scale feature extraction module, a feature fusion module, and a finely tuned adaptive decoder equipped with a visual-text alignment module. Notably, MC-Net possesses the capability to extract informative multiscale features, facilitated by the multilayer perceptron and transformer. We also introduce an adaptive gating mechanism during the decoding phase to ensure precise alignment between visual regions and their corresponding text descriptions. Empirical studies conducted on four publicly recognized cross-modal datasets unequivocally demonstrate the superior robustness and efficacy of MC-Net in comparison to contemporaneous RSIC methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
5  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
5  
ISSN:  
1753-8947  
eISSN:  
1753-8955  
入藏号:  
WOS:001125048100001  
Indexed Date:  
2024-01-06  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & Re, Wuhan, Peoples R China  
Institution Address:  
Huazhong Univ Sci & Technol, Sch Elect Informat & Commun, Wuhan, Peoples R China  
Institution Address:  
Univ Arkansas, Dept Geosci, Fayetteville, AR USA  
Institution Address:  
Sichuan Normal Univ, Sch Geog & Resources Sci, Chengdu, Sichuan, Peoples R China  
Institution Address:  
Univ Elect Sci & Technol, Sch Resources & Environm, Chengdu, Sichuan, Peoples R China  
Institution Address:  
Chengdu Univ Technol, Sch Geophys, Chengdu, Sichuan, Peoples R China  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & Re, Wuhan 430079, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 68. BCNet: integrating UNet and transformer for blood cell segmentation

摘要: Automatic segmentation of blood cells is crucial in medical diagnosis and research, significantly improving the accuracy and efficiency of diagnosing blood disorders. Traditional segmentation methods involving manual segmentation are time-consuming, labor-intensive, and prone to errors. In recent years, advancements in deep learning have provided new solutions for automated segmentation. This paper proposes BCNet, a blood cell segmentation algorithm combining UNet and Transformer. Specifically, BCNet utilizes UNet's Encoder-Decoder architecture as the backbone for extracting multi-scale features. A Spatial Reduction Transformer (SRT) Module is introduced for capturing long-range dependencies in the deepest downsampling layers to enhance sensitivity to local features. Additionally, coordinate attention is employed instead of skip connections for multi-scale feature fusion, enriching semantic information in deep features. Experimental results demonstrate that BCNet achieves superior Dice and IoU metrics compared to classical medical image segmentation models, facilitating automated analysis and medical diagnosis of blood cells.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001367067200001  
Indexed Date:  
2024-12-07  
Institution Address:  
Tiangong Univ, Sch Elect & Informat Engn, Tianjin Key Lab Optoelect Detect Technol & Syst, Tianjin 300387, Peoples R China  
Institution Address:  
Civil Aviat Univ China, Law Sch, Tianjin 300300, Peoples R China  
Institution Address:  
Natl Univ Def Technol, Coll Comp Sci, Changsha 410073, Peoples R China  
Institution:  
Institution:  
Institution:

### 69. 3D Point Cloud Object Detection Method Based on Multi-Scale Dynamic

摘要: Perception plays a crucial role in ensuring the safety and reliability of autonomous driving systems. However, the recognition and localization of small objects in complex scenarios still pose challenges. In this paper, we propose a point cloud object detection method based on dynamic sparse voxelization to enhance the detection performance of small objects. This method employs a specialized point cloud encoding network to learn and generate pseudo-images from point cloud features. The feature extraction part uses sliding windows and transformer-based methods. Furthermore, multi-scale feature fusion is performed to enhance the granularity of small object information. In this experiment, the term "small object" refers to objects such as cyclists and pedestrians, which have fewer pixels compared to vehicles with more pixels, as well as objects of poorer quality in terms of detection. The experimental results demonstrate that, compared to the PointPillars algorithm and other related algorithms on the KITTI public dataset, the proposed algorithm exhibits improved detection accuracy for cyclist and pedestrian target objects. In particular, there is notable improvement in the detection accuracy of objects in the moderate and hard quality categories, with an overall average increase in accuracy of about 5%.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001192903700001  
PubMed ID:  
38544067  
Indexed Date:  
2024-04-04  
Institution Address:  
Shanghai Inst Technol, Sch Comp Sci & Informat Technol, Shanghai 200235, Peoples R China  
Institution Address:  
Shenzhen Technol Univ, Coll Engn Phys, Shenzhen 518118, Peoples R China  
Institution:  
Institution:

### 70. Regional-to-Local Point-Voxel Transformer for Large-Scale Indoor 3D

摘要: Semantic segmentation of large-scale indoor 3D point cloud scenes is crucial for scene understanding but faces challenges in effectively modeling long-range dependencies and multi-scale features. In this paper, we present RegionPVT, a novel Regional-to-Local Point-Voxel Transformer that synergistically integrates voxel-based regional self-attention and window-based point-voxel self-attention for concurrent coarse-grained and fine-grained feature learning. The voxel-based regional branch focuses on capturing regional context and facilitating inter-window communication. The window-based point-voxel branch concentrates on local feature learning while integrating voxel-level information within each window. This unique design enables the model to jointly extract local details and regional structures efficiently and provides an effective and efficient solution for multi-scale feature fusion and a comprehensive understanding of 3D point clouds. Extensive experiments on S3DIS and ScanNet v2 datasets demonstrate that our RegionPVT achieves competitive or superior performance compared with state-of-the-art approaches, attaining mIoUs of 71.0% and 73.9% respectively, with significantly lower memory footprint.  
Web of Science 核心合集中的 "被引频次":  
5  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
5  
eISSN:  
2072-4292  
入藏号:  
WOS:001083064400001  
Indexed Date:  
2023-10-28  
Institution Address:  
Beijing Forestry Univ, Coll Sci, Beijing 100083, Peoples R China  
Institution:

### 71. HiFuse: Hierarchical multi-scale feature fusion network for medical

摘要: Effective fusion of global and local multi-scale features is crucial for medical image classification. Medical images have many noisy, scattered features, intra-class variations, and inter-class similarities. Many studies have shown that global and local features are helpful to reduce noise interference in medical images. It is difficult to capture the global features of images due to the fixed size of the receptive domain of the convolution kernel. Although the self-attention-based Transformer can model long-range dependencies, it has high computational complexity and lacks local inductive bias. In this paper, we propose a three-branch hierarchical multi-scale feature fusion network structure termed as HiFuse, which can fuse multi-scale global and local features without destroying the respective modeling, thus improving the classification accuracy of various medical images. There are two key characteristics: (i) a parallel hierarchical structure consisting of global and local feature blocks; (ii) an adaptive hierarchical feature fusion block (HFF block) and inverted residual multi-layer perceptron(IRMLP). The advantage of this network structure lies in that the resulting representation is semantically richer and the local features and global representations can be effectively extracted at different semantic scales. Our proposed model's ACC and F1 values reached 85.85% and 75.32% on the ISIC2018 dataset, 86.12% and 86.13% on the Kvasir dataset, 76.88% and 76.31% on the Covid-19 dataset, 92.31% and 88.81% on the esophageal cancer pathology dataset. The HiFuse model performs the best compared to other advanced models. Our code is open source and available from https://github.com/ huoxiangzuo/HiFuse.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
9  
在中国科学引文数据库中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
75  
被引频次合计:  
75  
ISSN:  
1746-8094  
eISSN:  
1746-8108  
入藏号:  
WOS:001088716700001  
Indexed Date:  
2023-11-05  
Institution Address:  
Xinjiang Univ, Sch Informat Sci & Engn, Urumqi 830000, Xinjiang, Peoples R China  
Institution Address:  
Xinjiang Univ, Xinjiang Key Lab Signal Detect & Proc, Urumqi 830000, Xinjiang, Peoples R China  
Institution Address:  
Xinjiang Med Univ, Dept Breast & Thyroid Surg, Affiliated Tumour Hosp, Urumqi 830011, Xinjiang, Peoples R China  
Institution Address:  
Xinjiang Med Univ, Xinjiang Canc Ctr, Key Lab Oncol, Affiliated Tumour Hosp, Urumqi 830011, Xinjiang, Peoples R China  
Institution Address:  
Cent South Univ, Big Data Inst, Changsha 410083, Hunan, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 72. Multi-scale object detection algorithm for recycled objects based on

摘要: An improved algorithm MR2-YOLOV5 based on YOLOv5 was proposed aiming at the problem that the traditional target detection algorithm did not consider the diversity of the target shape scale in the actual sorting scene and could not obtain the rotation angle information. Precise rotation angle detection was completed by adding angle prediction branches and introducing angle classification method of ring smooth label (CSL). The target detection layer was added to improve the detection ability of different scales of the model. Transformer attention mechanism was used at the end of the backbone network to give different weights to each channel and strengthen feature extraction. The feature graphs of different levels extracted from the backbone network were input into the BiFPN network structure to conduct multi-scale feature fusion. The experimental results showed that the mean average precision (mAP) of MR2-YOLOV5 on the self-made data set was 90.56%, which was 5.36% higher than that of YOLOv5s with only angle prediction branch. Categories and rotation angles can be recognized for objects such as occlusion, transparent and deformation. The detection time of single frame is 0.02-0.03 s, which meets the performance requirements of target detection algorithm for sorting scenes.  
摘要:  
针对传统目标检测算法未考虑实际分拣场景目标物形态尺度的多样性,无法获取旋转角度信息的问题,提出基于YOLOv5的改进算法MR2-YOLOv5.通过添加角度预测分支,引入环形平滑标签(CSL)角度分类方法,完成旋转角度精准检测.增加目标检测层用于提升模型不同尺度检测能力,在主干网络末端利用Transformer注意力机制对各通道赋予不同的权重,强化特征提取.利用主干网络提取到的不同层次特征图输入BiFPN网络结构中,开展多尺度特征融合.实验结果表明,MR2-YOLOv5在自制数据集上的均值平均精度(mAP)为90.56%,较仅添加角度预测分支的YOLOv5s基础网络提升5.36%;对于遮挡、透明、变形等目标物,均可以识别类别和旋转角度,图像单帧检测时间为0.02~0.03 s,满足分拣场景对目标检测算法的性能需求.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1008-973X  
入藏号:  
CSCD:7138049  
Indexed Date:  
2022-04-02  
Institution Address:  
浙江工业大学智能交通系统联合研究所, 杭州, 浙江 310014, 中国  
Institution Address:  
ITS Joint Research Institute, Zhejiang University of Technology, Hangzhou, Zhejiang 310014, China  
Institution:  
浙江工业大学智能交通系统联合研究所  
Institution:  
ITS Joint Research Institute, Zhejiang University of Technology

### 73. eViTBins: Edge-Enhanced Vision-Transformer Bins for Monocular Depth

摘要: Monocular depth estimation (MDE) remains a fundamental yet not well-solved problem in computer vision. Current wisdom of MDE often achieves blurred or even indistinct depth boundaries, degenerating the quality of vision-based intelligent transportation systems. This paper presents an edge-enhanced vision transformer bins network for monocular depth estimation, termed eViTBins. eViTBins has three core modules to predict monocular depth maps with exceptional smoothness, accuracy, and fidelity to scene structures and object edges. First, a multi-scale feature fusion module is proposed to circumvent the loss of depth information at various levels during depth regression. Second, an image-guided edge-enhancement module is proposed to accurately infer depth values around image boundaries. Third, a vision transformer-based depth discretization module is introduced to comprehend the global depth distribution. Meanwhile, unlike most MDE models that rely on high-performance GPUs, eViTBins is optimized for seamless deployment on edge devices, such as NVIDIA Jetson Nano and Google Coral SBC, making it ideal for real-time intelligent transportation systems applications. Extensive experimental evaluations corroborate the superiority of eViTBins over competing methods, notably in terms of preserving depth edges and global depth representations.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1524-9050  
eISSN:  
1558-0016  
入藏号:  
WOS:001342442300001  
Indexed Date:  
2024-11-06  
Institution Address:  
Nanjing Univ Aeronaut & Astronaut, Sch Comp Sci & Technol, Nanjing 211106, Peoples R China  
Institution Address:  
Sun Yat Sen Univ, Sch Geospatial Engn & Sci, Zhuhai 519082, Peoples R China  
Institution Address:  
Lingnan Univ, Sch Data Sci, Hong Kong, Peoples R China  
Institution Address:  
Hong Kong Metropolitan Univ, Sch Sci & Technol, Hong Kong, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 74. Lung Nodule Segmentation Algorithm With SMR-UNet

摘要: Accurate segmentation of lung nodules is of great significance for the early diagnosis of lung cancer. However, due to the diverse shapes and small sizes of lung nodules, lung nodule segmentation is a difficult task. In this paper, we propose an improved U-Net network called SMR-UNet, which integrates self-attention, multi-scale features and residual structures for lung nodule segmentation. The framework replaces the U-Net's convolutional units with residual units to ensure fast convergence, enhances the network's global modeling capability with Transformer, restores more detailed information with PixelShuffle, and enlarges the receptive field with a multi-scale feature fusion module before upsampling. The experiments show that on the LIDC dataset, the SMR-UNet achieves a Dice index of 0.9187 and an IoU of 0.8688, which are improved by 1.33% and 2.36% espectively compared to U-Net. When tested on lung nodules provided by the Department of Medical Imaging of the Fourth Affiliated Hospital of Guangxi Medical University, the SMR-UNet achieves a Dice index of 0.7785 and an IoU of 0.6541, which are improved by 4.22% and 4% espectively compared to U-Net.  
在 BIOSIS Citation Index 中的被引频次:  
1  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
10  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
11  
ISSN:  
2169-3536  
入藏号:  
WOS:000970909700001  
Indexed Date:  
2023-04-30  
Institution Address:  
Guangxi Med Univ, Dept Radiol, Affiliated Hosp 4, Liuzhou 545000, Guangxi, Peoples R China  
Institution Address:  
Guangxi Univ Sci & Technol, Sch Automat, Liuzhou 545006, Peoples R China  
Institution Address:  
Guangxi Univ Sci & Technol, Guangxi Collaborat Innovat Ctr Earthmoving Machine, Liuzhou 545006, Peoples R China  
Institution:  
Institution:  
Institution:

### 75. A Multi-scale Deformable Convolution Network Model for Text Recognition

摘要: Natural scene text recognition is one of the most challenging tasks in recent years. Compared with traditional document text, natural scene text has the characteristics of various shapes and different directions, so the accuracy of scene text recognition still needs to be improved. In order to locate the text region better and identify the text content more accurate, we present a multi-scale deformable convolution network model for text recognition. The initial image is irregularly corrected through the rectified network, and the ResNet with FPN structure is used as the backbone network to achieve multi-scale feature extraction. In addition, the feature fusion method of Add is adopted to reduce feature information losing and increase the strength of feature extraction in the text area. The deformable convolution block is introduced in the deep convolution to improve the deformation modeling ability of convolution and expand the receptive field. The prediction module adopts the Transformer and abandons the inherent pre and post attributes of RNN to realize parallel operation and solve the problem of path length between remote dependencies. In order to evaluate the effectiveness of the proposed method, we trained our model on two mixed data sets, MJSynth and SynthText, and tested it on some regular and irregular data sets. The experiment results demonstrate that this method performs well in irregular scene text recognition, especially in CUTE80.  
会议名称:  
13th International Conference on Graphics and Image Processing (ICGIP)  
会议日期:  
AUG 18-20, 2021  
会议地点:  
Yunnan Univ, Kunming, PEOPLES R CHINA  
会议赞助商:  
Yunnan Univ, Sch Informat Sci & Engn; Foshan Univ; Chang'an Univ; Xian Univ Technol; NW A & F Univ; Sichuan Univ; Ocean Univ China; Univ Portsmouth; China Soc Image & Graph; China Graph Soc  
会议主办方:  
Yunnan Univ  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0277-786X  
eISSN:  
1996-756X  
ISBN:  
978-1-5106-5043-5; 978-1-5106-5042-8  
入藏号:  
WOS:000796487100071  
Indexed Date:  
2022-05-27  
Institution Address:  
Jiangxi Normal Univ, Sch Comp & Informat Engn, Nanchang 330022, Jiangxi, Peoples R China  
Institution:

### 76. CAFS-Net: Context-Aware Feature Selection Network for Accurate and

摘要: Although deep learning-based methods have achieved remarkable performance in the industrial defect segmentation area, tiny defect segmentation in ultrahigh-resolution image scenarios still remains unexplored. Most of the existing methods utilize attention mechanisms and a sliding window strategy for tiny defect segmentation. However, this approach is not only computationally demanding but also prone to texture noise interference, likely stemming from a lack of global contextual understanding. To alleviate this challenge, we propose a context-aware feature selection network (CAFS-Net) which consists of a context-aware encoder network (CANet), a novel feature selection module (FSM), and a lightweight feature-fusing decoder. The CANet is constructed by low-level convolutional blocks and high-level transformer blocks to capture both local and global context information, thereby enhancing the discrimination ability between tiny defects and texture noise. The FSM includes a multilayer perceptron (MLP) classifier and a selector for selecting defective features of image blocks from the feature pyramid of image blocks based on the classification outcomes. Then, the selected defective features are fed to the lightweight feature-fusing decoder to perform multiscale feature fusion and obtain the segmentation masks of the defective image patches. Additionally, we propose a concentrated random cropping data augmentation strategy to address the class imbalance problem during training. We conducted extensive experiments on two defect segmentation datasets, including the compact camera module (CCM) defect segmentation dataset and the printed circuit board (PCB) defect segmentation dataset, to demonstrate the superiority and generalization performance of our proposed model. The results show that our CAFS-Net outperforms other state-of-the-art (SOTA) methods in both accuracy and efficiency.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001445061900011  
Indexed Date:  
2025-03-23  
Institution Address:  
Harbin Inst Technol, Res Inst Intelligent Control & Syst, Harbin 150001, Peoples R China  
Institution:

### 77. Multi-Scale Efficient Graph-Transformer for Whole Slide Image

摘要: The multi-scale information among the whole slide images (WSIs) is essential for cancer diagnosis. Although the existing multi-scale vision Transformer has shown its effectiveness for learning multi-scale image representation, it still cannot work well on the gigapixel WSIs due to their extremely large image sizes. To this end, we propose a novel Multi-scale Efficient Graph-Transformer (MEGT) framework for WSI classification. The key idea of MEGT is to adopt two independent efficient Graph-based Transformer (EGT) branches to process the low-resolution and high-resolution patch embeddings (i.e., tokens in a Transformer) of WSIs, respectively, and then fuse these tokens via a multi-scale feature fusion module (MFFM). Specifically, we design an EGT to efficiently learn the local-global information of patch tokens, which integrates the graph representation into Transformer to capture spatial-related information of WSIs. Meanwhile, we propose a novel MFFM to alleviate the semantic gap among different resolution patches during feature fusion, which creates a non-patch token for each branch as an agent to exchange information with another branch by cross-attention mechanism. In addition, to expedite network training, a new token pruning module is developed in EGT to reduce the redundant tokens. Extensive experiments on both TCGA-RCC and CAMELYON16 datasets demonstrate the effectiveness of the proposed MEGT.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
2168-2194  
eISSN:  
2168-2208  
入藏号:  
WOS:001147165700022  
PubMed ID:  
37725722  
Indexed Date:  
2024-02-03  
Institution Address:  
Shanghai Univ, Shanghai Inst Adv Commu & Data Sci, Key Lab Specialty Fiber Opt & Opt Access Networks, Sch Commun & Informat Engn,Joint Int Res Lab Spec, Shanghai 200444, Peoples R China  
Institution Address:  
Shanghai Univ, Sch Sci, Dept Math, Shanghai 200444, Peoples R China  
Institution:  
Institution:

### 78. HSD<SUP>2</SUP>Former: Hybrid-Scale Dual-Domain Transformer with

摘要: An unescapable trend of hyperspectral image (HSI) has been toward classification with high accuracy and splendid performance. In recent years, Transformers have made remarkable progress in the HSI classification task. However, Transformer-based methods still encounter two main challenges. First, they concentrate on extracting spectral information and are incapable of using spatial information to a great extent. Second, they lack the utilization of multiscale features and do not sufficiently combine the advantages of the Transformer's global feature extraction and multiscale feature extraction. To tackle these challenges, this article proposes a new solution named the hybrid-scale dual-domain Transformer with crisscrossed interaction (HSD2Former) for HSI classification. HSD2Former consists of three functional modules: dual-dimension multiscale convolutional embedding (D2MSCE), mixed domainFormer (MDFormer), and pyramid scale fusion block (PSFB). D2MSCE supersedes conventional patch embedding to generate spectral and spatial tokens at different scales, effectively enriching the diversity of spectral-spatial features. MDFormer is designed to facilitate self-enhancement and information interaction between the spectral domain and spatial domain, alleviating the heterogeneity of the spatial domain and spectral domain. PSFB introduces a straightforward fusion manner to achieve advanced semantic information for classification. Extensive experiments conducted on four datasets demonstrate the robustness and significance of HSD2Former. The classification evaluation indicators of OA, AA, and Kappa on four datasets almost exceed 98%, reaching state-of-the-art performance.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001377694100001  
Indexed Date:  
2024-12-21  
Institution Address:  
Chinese Acad Sci, Natl Key Lab Opt Field Manipulat Sci & Technol, Chengdu 610209, Peoples R China  
Institution Address:  
Chinese Acad Sci, Key Lab Opt Engn, Chengdu 610209, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Opt & Elect, Chengdu 610209, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 79. SCAD: A Siamese Cross-Attention Discrimination Network for Bitemporal

摘要: Building change detection (BCD) is crucial for urban construction and planning. The powerful discriminative ability of deep convolutions in deep learning-based BCD methods has considerably increased the accuracy and efficiency. However, dense and continuously distributed buildings contain a wide range of multi-scale features, which render current deep learning methods incapable of discriminating and incorporating multiple features effectively. In this work, we propose a Siamese cross-attention discrimination network (SCADNet) to identify complex information in bitemporal images and improve the change detection accuracy. Specifically, we first use the Siamese cross-attention (SCA) module to learn unchanged and changed feature information, combining multi-head cross-attention to improve the global validity of high-level semantic information. Second, we adapt a multi-scale feature fusion (MFF) module to integrate embedded tokens with context-rich channel transformer outputs. Then, upsampling is performed to fuse the extracted multi-scale information content to recover the original image information to the maximum extent. For information content with a large difference in contextual semantics, we perform filtering using a differential context discrimination (DCD) module, which can help the network to avoid pseudo-change occurrences. The experimental results show that the present SCADNet is able to achieve a significant change detection performance in terms of three public BCD datasets (LEVIR-CD, SYSU-CD, and WHU-CD). For these three datasets, we obtain F1 scores of 90.32%, 81.79%, and 88.62%, as well as OA values of 97.98%, 91.23%, and 98.88%, respectively.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
7  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
7  
eISSN:  
2072-4292  
入藏号:  
WOS:000902849300001  
Indexed Date:  
2023-01-08  
Institution Address:  
Hubei Univ Technol, Sch Comp Sci, Wuhan 430068, Peoples R China  
Institution Address:  
Wuhan Univ, Inst Technol Sci, Wuhan 430072, Peoples R China  
Institution Address:  
Naval Engn Univ, Sch Weap Engn, Wuhan 430032, Peoples R China  
Institution Address:  
Wuhan Univ, State Key Lab Informat Engn Surveying Mapping & Re, Wuhan 430072, Peoples R China  
Institution Address:  
Wuchang Shouyi Univ, Sch Informat Sci & Engn, Wuhan 430064, Peoples R China  
Institution Address:  
Air Force Res Acad, Beijing 10085, Peoples R China  
Institution:  
Institution:  
Institution:  
Naval Engn Univ  
Institution:  
Institution:  
Institution:  
Air Force Res Acad

### 80. Remote Sensing Image Change Detection Based on Deep Multi-Scale

摘要: Change detection is a technique that can observe changes in the surface of the earth dynamically. It is one of the most significant tasks in remote sensing image processing. In the past few years, with the ability of extracting rich deep image features, the deep learning techniques have gained popularity in the field of change detection. In order to obtain obvious image change information, the attention mechanism is added in the decoder and output stage in many deep learning-based methods. Many of these approaches neglect to upgrade the ability of the encoders and the feature extractors to extract the representational features. To resolve this problem, this study proposes a deep multi-scale multi-attention siamese transformer network. A special contextual attention module combining a convolution and self-attention module is introduced into the siamese feature extractor to enhance the global representation ability. A lightly efficient channel attention block is added in the siamese feature extractor to obtain the information interaction among different channels. Furthermore, a multi-scale feature fusion module is proposed to fuse the features from different stages of the siamese feature extractor, and it can detect objects of different sizes and irregularities. To increase the accuracy of the proposed approach, the transformer module is utilized to model the long-range context in two-phase images. The experimental results on the LEVIR-CD and the CCD datasets show the effectiveness of the proposed network.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
35  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
1  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
37  
eISSN:  
2072-4292  
入藏号:  
WOS:000932577100001  
Indexed Date:  
2023-03-16  
Institution Address:  
Xidian Univ, Sch Artificial Intelligence, Key Lab Intelligent Percept & Image Understanding, Minist Educ, Xian 710071, Peoples R China  
Institution Address:  
Xidian Univ, Sch Elect Engn, Xian 710071, Peoples R China  
Institution:  
Institution:

### 81. Gateinst: instance segmentation with multi-scale gated-enhanced queries

摘要: Recently, a popular query-based end-to-end framework has been used for instance segmentation. However, queries update based on individual layers or scales of feature maps at each stage of Transformer decoding, which makes queries unable to gather sufficient multi-scale feature information. Therefore, querying these features may result in inconsistent information due to disparities among feature maps and leading to erroneous updates. This study proposes a new network called GateInst, which employs a dual-path auto-select mechanism based on gate structures to overcome these issues. Firstly, we design a block-wise multi-scale feature fusion module that combines features of different scales while maintaining low computational cost. Secondly, we introduce the gated-enhanced queries Transformer decoder that utilizes a gating mechanism to filter and merge the queries generated at different stages to compensate for the inaccuracies in updating queries. GateInst addresses the issue of insufficient feature information and compensates for the problem of cumulative errors in queries. Experiments have shown that GateInst achieves significant gains of 8.4 AP, 5.5 AP50\documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{mathrsfs} \usepackage{upgreek} \setlength{\oddsidemargin}{-69pt} \begin{document}$$AP\_{50}$$\end{document} over Mask2Former on the self-collected Tree Species Instance Dataset and performs well compared to non-Mask2Former-like and Mask2Former-like networks on self-collected and public COCO datasets, with only a tiny amount of additional computational cost and fast convergence. Code and models are available at https://github.com/FAFU-IMLab/GateInst.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0942-4962  
eISSN:  
1432-1882  
入藏号:  
WOS:001294481100003  
Indexed Date:  
2024-08-24  
Institution Address:  
Fujian Agr & Forestry Univ, Coll Comp & Informat Sci, Fuzhou 350002, Fujian, Peoples R China  
Institution:

### 82. Dual attention and channel transformer based generative adversarial

摘要: Artworks are treasures of valuable cultural and historical heritage. Artworks get damaged due to environmental and other factors. The artificial intelligence-based restoration of digitized artwork images can guide the artists in physically restoring the damaged artworks. Previous methods have not been able to restore artwork images well. This paper proposes a dual (spatial and channel) attention and channel transformer-based generative adversarial network to restore damaged artwork images digitally. The proposed generative adversarial network has spatial and channel attention layers in the encoder part of the generator and a channel transformer between skip connections from the encoder to the decoder part of the generator. Spatial and channel attention helps learn inter-spatial and inter-channel global relationships among image features. Channel transformer ensures multiscale feature fusion and reduces the semantic gap between encoder and decoder layer features. Moreover, the proposed network has been trained using a linear combination of perceptual, adversarial, and structured similarity index measure loss, which helps better train the network. Further, the proposed network has been validated on two different datasets, and the results indicate that the proposed method outperforms state-of-the-art artwork restoration methods.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
5  
在 BIOSIS Citation Index 中的被引频次:  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
5  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001166366700001  
Indexed Date:  
2024-03-06  
Institution Address:  
Chandigarh Coll Engn & Technol, Dept Comp Sci & Engn, Chandigarh, India  
Institution Address:  
CCET Degree Wing, Sect 26, Chandigarh 160019, India  
Institution:  
Chandigarh Coll Engn & Technol  
Institution:  
CCET Degree Wing

### 83. Feature Aggregation and Compensation-Based Domain Adaptation Framework

摘要: Synthetic aperture radar (SAR) target recognition plays an indispensable role in interpreting SAR images. However, differences in radar parameters (including factors such as imaging modes and imaging angles) often lead to resolution differences between training and test data, posing challenges for existing methods in recognizing SAR targets under cross-resolution conditions. To address this issue, this article proposes a domain adaptation (DA) framework based on feature aggregation and compensation (FAC) for cross-resolution target recognition in SAR imagery. Initially, we employ a unique local vision transformer (LocalViT) to establish global and local adversarial networks that capture invariant features under cross-resolution conditions. Following this, we design a multiscale feature fusion module (MSFFM) to capture multiscale semantic features of targets at different resolutions. Subsequently, we propose a novel class feature aggregation module (CFAM) to map targets of varying resolutions onto the unit sphere, thereby aggregating features of samples from the same class and distinguishing those of samples from different classes. Finally, we narrow down the difference in frequency-domain information of targets at different resolutions by developing a resolution semantic compensation module (RSCM). This module compensates for the semantic feature about resolution during target recognition across varying resolutions by converting high- and low-frequency information. The experimental results on three SAR datasets (OpenSARShip, FUSAR-Ship, and SRSDD-v1.0) confirm that our method outperforms the state-of-the-art (SOTA) DA methods, with an increase of 2.06%, 1.96%, and 1.61% in three sets of cross-resolution scenes, respectively.  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001383067500001  
Indexed Date:  
2024-12-29  
Institution Address:  
Univ Elect Sci & Technol China, Sch Informat & Commun Engn, Chengdu 611731, Peoples R China  
Institution:

### 84. Position-sensitive Transformer aerial image object detection model

摘要: Addressing the challenge of detecting numerous small objects in UAV-captured aerial images, this paper introduces the Position-Sensitive Transformer Target Detection (PS-TOD) model. Initially, it presents a multi-scale feature fusion (MSFF) module incorporating a Positional Channel Embedded 3D Attention (PCE3DA) mechanism. PCE3DA leverages the interplay between spatial and channel data to generate 3D attention, enhancing feature representation in areas of interest. This foundation supports a bottomup, cross-layer MSFF approach, augmenting the semantic richness of combined features. Subsequently, it proposes a novel Position-Sensitive Self-Attention (PSSA) mechanism, leading to the development of a position-sensitive Transformer encoder-decoder. This innovation heightens the model's sensitivity to target positioning, facilitating the capture of long-term dependencies within the image's global context. Comparative tests using the VisDrone dataset reveal that the PS-TOD model attains an Average Precision (AP) of 28.8%, marking a 4.1% enhancement over the baseline model (DETR). Furthermore, it demonstrates precise object detection in UAV aerial imagery against complex backdrops, significantly boosting the detection accuracy of small targets.  
摘要:  
针对无人机视角下航拍图像小目标多且检测困难的问题,提出了一个位置敏感Transformer目标检测(PS-TOD)模型。设计了一个基于位置通道嵌入三维注意力(PCE3DA)的多尺度特征融合(MSFF)模块,即PCE3DA利用空间与通道信息的相互依赖关系生成三维注意力,用于加强模型对兴趣区域的特征表达能力,且基于它构造了一个自底向上的跨层MSFF方案,使得融合后的特征语义信息更加丰富;然后,设计了一种新的位置敏感自注意力(PSSA)机制,且以此构造位置敏感Transformer编-解码器,使模型在捕获图像全局上下文信息的长期依赖关系时,也可提高模型对目标的位置敏感能力。基于无人机航拍数据集VisDrone的对比实验结果表明,提出模型的AP达到28.8%,与基线模型(DETR)相比提高了4.1%。该模型在复杂背景下能对无人机航拍图像进行精确的目标检测,且改善小目标的检测效果。  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1004-924X  
入藏号:  
CSCD:7669310  
Indexed Date:  
2024-05-17  
Institution Address:  
西安邮电大学通信与信息工程学院, 西安, 陕西 710121, 中国  
Institution Address:  
College of communication and information engineering, Xi′an University of Posts and Telecommunication, Xi'an, Shaanxi 710121, China  
Institution:  
西安邮电大学通信与信息工程学院  
Institution:  
College of communication and information engineering, Xi′an University of Posts and Telecommunication

### 85. Infrared and Visible Images Fusion Method Based on Multi-Scale Features

摘要: To address the challenges of detail loss and the imbalance between visual detail features and infrared(IR) target features in fused infrared and visible images,this study proposes a fusion method combining multiscale feature fusion and efficient multi-head self-attention (EMSA).The method includes several key steps.1) Multiscale coding network:It utilizes a multiscale coding network to extract multilevel features,enhancing the descriptive capability of the scene.2) Fusion strategy:It combines transformer-based EMSA with dense residual blocks to address the imbalance between local details and overall structure in the fusion process.3) Nested-connection based decoding network:It takes the multilevel fusion map and feeds it into a nested-connection based decoding network to reconstruct the fused result,emphasizing prominent IR targets and rich scene details.Extensive experiments on the TNO and M3FD public datasets demonstrate the efficacy of the proposed method.It achieves superior results in both quantitative metrics and visual comparisons.Specifically,the proposed method excels in targeted detection tasks,demonstrating state-of-the-art performance.This approach not only enhances the fusion quality by effectively preserving detailed information and balancing visual and IR features but also establishes a benchmark in the field of infrared and visible image fusion.  
摘要:  
针对红外与可见光图像融合容易出现细节丢失,且现有的融合策略难以平衡视觉细节特征和红外目标特征等问题,提出一种基于多尺度特征融合与高效多头自注意力相结合的红外与可见光图像融合方法。首先,为提高目标与场景的描述能力,采用了多尺度编码网络提取源图像不同尺度的特征;其次,提出了基于Transformer的多头转置注意力结合残差密集块的融合策略以平衡融合细节与整体结构;最后,将多尺度特征融合图输入基于巢式连接的解码网络,重建具有显著红外目标和丰富细节信息的融合图像。基于TNO与M3FD公开数据集与7种经典融合方法进行实验,结果表明,本文方法在视觉效果与量化评价指标上表现更佳,生成的融合图像在目标检测任务上取得更好的效果。  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1001-8891  
入藏号:  
CSCD:7765424  
Indexed Date:  
2024-09-21  
Institution Address:  
西南科技大学信息工程学院;;特殊环境机器人技术四川省重点实验室, ;;特殊环境机器人技术四川省重点实验室, 绵阳;;绵阳, 四川;;四川 621010;;621010, 中国  
Institution Address:  
中国电信股份有限公司成都分公司, 成都, 四川 610066, 中国  
Institution Address:  
中国空气动力研究与发展中心空天技术研究所, 绵阳, 四川 621006, 中国  
Institution Address:  
西南科技大学信息工程学院, 绵阳, 四川 621010, 中国  
Institution Address:  
School of Information Engineering,Southwest University of Science and Technology;;Sichuan Key Laboratory of Special Environmental Robotics, ;;Sichuan Key Laboratory of Special Environmental Robotics, Mianyang;;Mianyang, ;; 621010;;621010  
Institution Address:  
China Telecom Corporation,Chengdu Branch, Chengdu, Sichuan 610066, China  
Institution Address:  
Institute of Aerospace Technology,China Aerodynamics Research and Development Center, Mianyang, Sichuan 621006, China  
Institution Address:  
School of Information Engineering,Southwest University of Science and Technology, Mianyang, Sichuan 621010, China  
Institution:  
西南科技大学信息工程学院;;特殊环境机器人技术四川省重点实验室  
Institution:  
中国电信股份有限公司成都分公司  
Institution:  
中国空气动力研究与发展中心空天技术研究所  
Institution:  
西南科技大学信息工程学院  
Institution:  
School of Information Engineering,Southwest University of Science and Technology;;Sichuan Key Laboratory of Special Environmental Robotics  
Institution:  
China Telecom Corporation,Chengdu Branch  
Institution:  
Institute of Aerospace Technology,China Aerodynamics Research and Development Center  
Institution:  
School of Information Engineering,Southwest University of Science and Technology

### 86. Two-Stage Self-Supervised Contrastive Learning Aided Transformer for

摘要: The availability of large, high-quality annotated datasets in the medical domain poses a substantial challenge in segmentation tasks. To mitigate the reliance on annotated training data, self-supervised pre-training strategies have emerged, particularly employing contrastive learning methods on dense pixel-level representations. In this work, we proposed to capitalize on intrinsic anatomical similarities within medical image data and develop a semantic segmentation framework through a self-supervised fusion network, where the availability of annotated volumes is limited. In a unified training phase, we combine segmentation loss with contrastive loss, enhancing the distinction between significant anatomical regions that adhere to the available annotations. To further improve the segmentation performance, we introduce an efficient parallel transformer module that leverages Multiview multiscale feature fusion and depth-wise features. The proposed transformer architecture, based on multiple encoders, is trained in a self-supervised manner using contrastive loss. Initially, the transformer is trained using an unlabeled dataset. We then fine-tune one encoder using data from the first stage and another encoder using a small set of annotated segmentation masks. These encoder features are subsequently concatenated for the purpose of brain tumor segmentation. The multiencoder-based transformer model yields significantly better outcomes across three medical image segmentation tasks. We validated our proposed solution by fusing images across diverse medical image segmentation challenge datasets, demonstrating its efficacy by outperforming state-of-the-art methodologies.  
Web of Science 核心合集中的 "被引频次":  
4  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
eISSN:  
2168-2208  
入藏号:  
MEDLINE:38090821  
PubMed ID:  
38090821  
Indexed Date:  
2023-12-16

### 87. PR-Deformable DETR: DETR for Remote Sensing Object Detection

摘要: Identifying objects in remote sensing images remains a critical challenge. However, remote sensing images typically encompass numerous small objects, significant variations in object sizes, and a dispersed distribution of objects, all of which pose challenges to the performance of existing object detectors. We present PR-Deformable DEtection Transformer (DETR), a novel model for remote sensing object detection to address these challenges. First, we introduce the tridirectional adaptive feature fusion pyramid network (TAFFPN) feature pyramid module to adaptively fuse data from diverse feature map layers, thereby enhancing the model's multiscale representation capability. Second, we propose the Res-Deformable Encoder, which integrates deformable encoders across different input scales via residual connections, generating feature vectors that capture rich semantic information of remote sensing objects. Last, we introduce the dynamic reference point module (DRPM) Decoder, which leverages 4-D reference points enriched with high-level (HL) feature priors to strengthen the model's object localization capabilities. Experimental results demonstrate that PR-Deformable DETR achieves state-of-the-art remote sensing object detection accuracy, achieving 88.3% mean average precision (mAP) on the NWPU VHR-10 dataset and 95.1% mAP on the RSOD dataset, with a corresponding 16% reduction in GFLOPs. These results satisfy the performance standards required for remote sensing object detection tasks.  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1545-598X  
eISSN:  
1558-0571  
入藏号:  
WOS:001349449300005  
Indexed Date:  
2024-11-25  
Institution Address:  
Wuhan Univ Technol, Sch Automat, Wuhan 430070, Peoples R China  
Institution:

### 88. Transformer-Enhanced Retinal Vessel Segmentation for Diabetic

摘要: Eye health has become a significant concern in recent years, given the rising prevalence of visual impairment resulting from various eye disorders and related factors. Global surveys suggest that approximately 2.2 billion individuals are visually impaired, with at least 1 billion affected by treatable diseases or ailments. Early detection, treatment, and screening for fundus diseases are crucial in addressing these challenges. In this study, we propose a novel segmentation model for retinal vascular delineation aimed at diagnosing diabetic retinopathy. The model integrates CBAM (Channel-Attention and Spatial-Attention) for enhanced feature representation, JPU (Joint Pyramid Upsampling) for multi-scale feature fusion, and transformer blocks for contextual understanding. Leveraging deep-learning techniques, our proposed model outperforms existing approaches in retinal vascular segmentation, like achieving a Mean IOU of 0.8047, Recall of 0.7254, Precision of 0.8492, F1 Score of 0.7824, and Specificity of 0.9892 for CHASEDB1 dataset. Extensive evaluations on benchmark datasets demonstrate its efficacy, highlighting its potential for automated diabetic retinopathy screening.  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
在 BIOSIS Citation Index 中的被引频次:  
1  
被引频次合计:  
4  
eISSN:  
2076-3417  
入藏号:  
WOS:001366897700001  
Indexed Date:  
2024-12-07  
Institution Address:  
Elect & Telecommun Res Inst, Daejeon 34129, South Korea  
Institution Address:  
Jeonbuk Natl Univ, Dept Elect & Informat Engn, Jeonju 54896, South Korea  
Institution Address:  
Jeonbuk Natl Univ, Adv Elect & Informat Res Ctr, Jeonju 54896, South Korea  
Institution:  
Institution:  
Institution:

### 89. TMFN: a text-based multimodal fusion network with multi-scale feature

摘要: Multimodal sentiment analysis (MSA) is crucial in human-computer interaction. Current methods use simple sub-models for feature extraction, neglecting multi-scale features and the complexity of emotions. Text, visual, and audio each have unique characteristics in MSA, with text often providing more emotional cues due to its rich semantics. However, current approaches treat modalities equally, not maximizing text's advantages. To solve these problems, we propose a novel method named a text-based multimodal fusion network with multi-scale feature extraction and unsupervised contrastive learning (TMFN). Firstly, we propose an innovative pyramid-structured multi-scale feature extraction method, which captures the multi-scale features of modal data through convolution kernels of different sizes and strengthens key features through channel attention mechanism. Second, we design a text-based multimodal feature fusion module, which consists of a text gating unit (TGU) and a text-based channel-wise attention transformer (TCAT). TGU is responsible for guiding and regulating the fusion process of other modal information, while TCAT improves the model's ability to capture the relationship between features of different modalities and achieves effective feature interaction. Finally, to further optimize the representation of fused features, we introduce unsupervised contrastive learning to deeply explore the intrinsic connection between multi-scale features and fused features. Experimental results show that our proposed model outperforms the state-of-the-art models in MSA on two benchmark datasets.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2199-4536  
eISSN:  
2198-6053  
入藏号:  
WOS:001391697300004  
Indexed Date:  
2025-01-13  
Institution Address:  
Chongqing Univ Technol, Coll Comp Sci & Engn, 69,Hongguang Ave, Chongqing 400054, Peoples R China  
Institution Address:  
Chongqing Univ Technol, Coll Artificial Intelligence, Chongqing 401135, Peoples R China  
Institution:  
Institution:

### 90. Mix-layers semantic extraction and multi-scale aggregation transformer

摘要: Recently, a number of vision transformer models for semantic segmentation have been proposed, with the majority of these achieving impressive results. However, they lack the ability to exploit the intrinsic position and channel features of the image and are less capable of multi-scale feature fusion. This paper presents a semantic segmentation method that successfully combines attention and multiscale representation, thereby enhancing performance and efficiency. This represents a significant advancement in the field. Multi-layers semantic extraction and multi-scale aggregation transformer decoder (MEMAFormer) is proposed, which consists of two components: mix-layers dual channel semantic extraction module (MDCE) and semantic aggregation pyramid pooling module (SAPPM). The MDCE incorporates a multi-layers cross attention module (MCAM) and an efficient channel attention module (ECAM). In MCAM, horizontal connections between encoder and decoder stages are employed as feature queries for the attention module. The hierarchical feature maps derived from different encoder and decoder stages are integrated into key and value. To address long-term dependencies, ECAM selectively emphasizes interdependent channel feature maps by integrating relevant features across all channels. The adaptability of the feature maps is reduced by pyramid pooling, which reduces the amount of computation without compromising performance. SAPPM is comprised of several distinct pooled kernels that extract context with a deeper flow of information, forming a multi-scale feature by integrating various feature sizes. The MEMAFormer-B0 model demonstrates superior performance compared to SegFormer-B0, exhibiting gains of 4.8%, 4.0% and 3.5% on the ADE20K, Cityscapes and COCO-stuff datasets, respectively.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2199-4536  
eISSN:  
2198-6053  
入藏号:  
WOS:001355112100005  
Indexed Date:  
2024-11-23  
Institution Address:  
Shandong Normal Univ, Sch Phys & Elect, Jinan, Shandong, Peoples R China  
Institution:

### 91. DV-DETR: Improved UAV Aerial Small Target Detection Algorithm Based on

摘要: For drone-based detection tasks, accurately identifying small-scale targets like people, bicycles, and pedestrians remains a key challenge. In this paper, we propose DV-DETR, an improved detection model based on the Real-Time Detection Transformer (RT-DETR), specifically optimized for small target detection in high-density scenes. To achieve this, we introduce three main enhancements: (1) ResNet18 as the backbone network to improve feature extraction and reduce model complexity; (2) the integration of recalibration attention units and deformable attention mechanisms in the neck network to enhance multi-scale feature fusion and improve localization accuracy; and (3) the use of the Focaler-IoU loss function to better handle the imbalanced distribution of target scales and focus on challenging samples. Experimental results on the VisDrone2019 dataset show that DV-DETR achieves an mAP@0.5 of 50.1%, a 1.7% improvement over the baseline model, while increasing detection speed from 75 FPS to 90 FPS, meeting real-time processing requirements. These improvements not only enhance the model's accuracy and efficiency but also provide practical significance in complex, high-density urban environments, supporting real-world applications in UAV-based surveillance and monitoring tasks.  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001366098600001  
PubMed ID:  
39599152  
Indexed Date:  
2024-12-05  
Institution Address:  
Shanghai Univ Engn Sci, Sch Elect & Elect Engn, Shanghai 201620, Peoples R China  
Institution:

### 92. Why Not Both? An Attention-Guided Transformer with Pixel-Related

摘要: Transformer-based encoder-decoder networks for face super-resolution (FSR) have achieved promising success in delivering stunningly clear and detailed facial images by capturing local and global dependencies. However, these methods have certain limitations. Specifically, the deconvolution in upsampling layers neglects the relationship between adjacent pixels, which is crucial in facial structure reconstruction. Additionally, raw feature maps are fed to the transformer blocks directly without mining their potential feature information, resulting in suboptimal face images. To circumvent these problems, we propose an attention-guided transformer with pixel-related deconvolution network for FSR. Firstly, we devise a novel Attention-Guided Transformer Module (AGTM), which is composed of an Attention-Guiding Block (AGB) and a Channel-wise Multi-head Transformer Block (CMTB). AGTM at the top of the encoder-decoder network (AGTM-T) promotes both local facial details and global facial structures, while AGTM at the bottleneck side (AGTM-B) optimizes the encoded features. Secondly, a Pixel-Related Deconvolution (PRD) layer is specially designed to establish direct relationships among adjacent pixels in the upsampling process. Lastly, we develop a Multi-scale Feature Fusion Module (MFFM) to fuse multi-scale features for better network flexibility and reconstruction results. Quantitative and qualitative experimental results on various datasets demonstrate that the proposed method outperforms other state-of-the-art FSR methods.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001219982200001  
Indexed Date:  
2024-05-21  
Institution Address:  
Xi An Jiao Tong Univ, Sch Elect & Informat Engn, Xian 710049, Peoples R China  
Institution:

### 93. Feature Maps Need More Attention: A Spatial-Channel Mutual

摘要: Recently, transformer-based face super-resolution (FSR) approaches have achieved promising success in restoring degraded facial details due to their high capability for capturing both local and global dependencies. However, while existing methods focus on introducing sophisticated structures, they neglect the potential feature map information, limiting FSR performance. To circumvent this problem, we carefully design a pair of guiding blocks to dig for possible feature map information to enhance features before feeding them to transformer blocks. Relying on the guiding blocks, we propose a spatial-channel mutual attention-guided transformer network for FSR, for which the backbone architecture is a multi-scale connected encoder-decoder. Specifically, we devise a novel Spatial-Channel Mutual Attention-guided Transformer Module (SCATM), which is composed of a Spatial-Channel Mutual Attention Guiding Block (SCAGB) and a Channel-wise Multi-head Transformer Block (CMTB). SCATM on the top layer (SCATM-T) aims to promote both local facial details and global facial structures, while SCATM on the bottom layer (SCATM-B) seeks to optimize the encoded features. Considering that different scale features are complementary, we further develop a Multi-scale Feature Fusion Module (MFFM), which fuses features from different scales for better restoration performance. Quantitative and qualitative experimental results on various datasets indicate that the proposed method outperforms other state-of-the-art FSR methods.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2076-3417  
入藏号:  
WOS:001234314400001  
Indexed Date:  
2024-06-05  
Institution Address:  
Xi An Jiao Tong Univ, Sch Elect & Informat Engn, Xian 710049, Peoples R China  
Institution:

### 94. MFINet: Multi-Scale Feature Interaction Network for Change Detection of

摘要: Change detection is widely used in the field of building monitoring. In recent years, the progress of remote sensing image technology has provided high-resolution data. However, unlike other tasks, change detection focuses on the difference between dual-input images, so the interaction between bi-temporal features is crucial. However, the existing methods have not fully tapped the potential of multi-scale bi-temporal features to interact layer by layer. Therefore, this paper proposes a multi-scale feature interaction network (MFINet). The network realizes the information interaction of multi-temporal images by inserting a bi-temporal feature interaction layer (BFIL) between backbone networks at the same level, guides the attention to focus on the difference region, and suppresses the interference. At the same time, a double temporal feature fusion layer (BFFL) is used at the end of the coding layer to extract subtle difference features. By introducing the transformer decoding layer and improving the recovery effect of the feature size, the ability of the network to accurately capture the details and contour information of the building is further improved. The F1 of our model on the public dataset LEVIR-CD reaches 90.12%, which shows better accuracy and generalization performance than many state-of-the-art change detection models.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
12  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
12  
eISSN:  
2072-4292  
入藏号:  
WOS:001201014300001  
Indexed Date:  
2024-04-18  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Collaborat Innovat Ctr Atmospher Environm & Equipm, Nanjing 210044, Peoples R China  
Institution Address:  
Nanjing Forestry Univ, Coll Informat Sci & Technol, Nanjing 210000, Peoples R China  
Institution:  
Institution:

### 95. Precise Facial Landmark Detection by Reference Heatmap Transformer

摘要: Most facial landmark detection methods predict landmarks by mapping the input facial appearance features to landmark heatmaps and have achieved promising results. However, when the face image is suffering from large poses, heavy occlusions and complicated illuminations, they cannot learn discriminative feature representations and effective facial shape constraints, nor can they accurately predict the value of each element in the landmark heatmap, limiting their detection accuracy. To address this problem, we propose a novel Reference Heatmap Transformer (RHT) by introducing reference heatmap information for more precise facial landmark detection. The proposed RHT consists of a Soft Transformation Module (STM) and a Hard Transformation Module (HTM), which can cooperate with each other to encourage the accurate transformation of the reference heatmap information and facial shape constraints. Then, a Multi-Scale Feature Fusion Module (MSFFM) is proposed to fuse the transformed heatmap features and the semantic features learned from the original face images to enhance feature representations for producing more accurate target heatmaps. To the best of our knowledge, this is the first study to explore how to enhance facial landmark detection by transforming the reference heatmap information. The experimental results from challenging benchmark datasets demonstrate that our proposed method outperforms the state-of-the-art methods in the literature.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
3  
在中国科学引文数据库中的被引频次:  
2  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
40  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
41  
ISSN:  
1057-7149  
eISSN:  
1941-0042  
入藏号:  
WOS:000964720100001  
PubMed ID:  
37030695  
Indexed Date:  
2023-05-09  
Institution Address:  
Zhongnan Univ Econ & Law, Sch Informat & Safety Engn, Wuhan 430073, Peoples R China  
Institution Address:  
Singapore Univ Technol & Design, Informat Syst Technol & Design Pillar, Singapore 487372, Singapore  
Institution Address:  
Shenzhen Univ, Coll Comp Sci & Software Engn, Shenzhen 518060, Peoples R China  
Institution Address:  
Robot Soc, Shenzhen Inst Artificial Intelligence, Shenzhen 518060, Peoples R China  
Institution Address:  
China Three Gorges Univ, Coll Comp & Informat Technol, Yichang 443002, Hubei, Peoples R China  
Institution Address:  
Zhongnan Univ Econ & Law, Sch Informat & Safety Engn, Wuhan 430073, Peoples R China  
Institution Address:  
Yunnan Univ, Sch Informat Sci & Engn, Kunming 650091, Yunnan, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 96. CG-SOLOv2: Enhanced instance segmentation for coal-gangue with novel

摘要: Computer-vision-based coal preparation has attracted widespread attention for its green and intelligent characteristics. However, existing studies predominantly focuses on coal-gangue object detection. They fail to adequately capture the shape and size of instances, thereby impairing the performance of downstream sorting mechanisms. Therefore, we developed an effective and efficient coal-gangue instance segmentation algorithm, named CG-SOLOv2. Specifically, we develop a novel feature extraction network, termed as CoNexTF. It merges convolutional local modeling with Transformer-based global modeling to enhance multi-scale feature extraction. Moreover, we develop Global Attention Feature Pyramid Network (GAFPN) as a new neck module. It enhances feature interaction across different scales, facilitating superior feature fusion. Experimental results demonstrate that our proposed model outperforms several state-of-the-art baseline models with a 1.36 %-25.07 % increase in Average Precision (AP) on laboratory datasets and 2.24 %-12.23 % on industrial datasets, respectively. These findings suggest promising prospects for its application in vision-based coal-gangue sorting practices.  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0032-5910  
eISSN:  
1873-328X  
入藏号:  
WOS:001394021200001  
Indexed Date:  
2025-01-17  
Institution Address:  
China Univ Min & Technol, Sch Chem Engn & Technol, Key Lab Coal Proc & Efficient Utilizat, Minist Educ, Xuzhou 221116, Jiangsu, Peoples R China  
Institution Address:  
Xinjiang Inst Engn, Coll Min & Geol, Urumqi 830023, Xinjiang, Peoples R China  
Institution Address:  
Univ Waterloo, Dept Mech & Mechatron Engn, 200 Univ Ave West, Waterloo, ON N2L 3G1, Canada  
Institution Address:  
Lakes Environm Res Inc, 170 Columbia St W Suite 1, Waterloo, ON N2L 3L3, Canada  
Institution:  
Institution:  
Institution:  
Institution:  
Lakes Environm Res Inc

### 97. Transformer tracking with multi-scale dual-attention

摘要: Transformer-based trackers greatly improve tracking success rate and precision rate. Attention mechanism in Transformer can fully explore the context information across successive frames. Nevertheless, it ignores the equally important local information and structured spatial information. And irrelevant regions may also affect the template features and search region features. In this work, a multi-scale feature fusion network is designed with box attention and instance attention in Encoder-Decoder architecture based on Transformer. After extracting features, the local information and structured spatial information is learnt by multi-scale box attention, and the global context information is explored by instance attention. Box attention samples grid features from the region of interest. Therefore, it effectively focuses on the region of interest (ROI) and avoids the influence of irrelevant regions in feature extraction. At the same time, instance attention can also pay attention to the context information across successive frames, and avoid falling into local optimum. The long-range feature dependencies are learned in this stage. Extensive experiments are conducted on six challenging tracking datasets to demonstrate the superiority of the proposed tracker MDTT, including UAV123, GOT-10k, LaSOT, VOT2018, TrackingNet, and NfS. In particular, the proposed tracker achieves AUC score of 64.7% on LaSOT, 78.1% on TrackingNet and precision score of 89.2% on UAV123, which outperforms the baseline and most recent advanced trackers.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
2199-4536  
eISSN:  
2198-6053  
入藏号:  
WOS:000964394500002  
Indexed Date:  
2023-04-28  
Institution Address:  
Nanchang Inst Technol, Sch Informat Engn, Nanchang 330029, Peoples R China  
Institution Address:  
NSFOCUS Technol Grp Co Ltd, Beijing 100089, Peoples R China  
Institution:  
Institution:  
NSFOCUS Technol Grp Co Ltd

### 98. A method for named entity recognition in social media texts with

摘要: Social media data are characterized by significant noise and non-standardization, thereby posing challenges for existing methods in recognizing named entities owing to the entity sparsity and insufficient semantic richness. Thus, to deal with these issues, this study proposes SEMFF-NER, a named entity recognition (NER) method in social media texts that integrates multi-scale features and syntactic information. First, global features are extracted using a Transformer-based encoder (XLNET) with embedded dependency syntactic relations to enhance semantic representation. Next, sliding windows of different lengths capture local features, which are input into a bi-directional long short-term memory (BiLSTM) to capture multi-level local features. Subsequently, the fusion-attention mechanism effectively integrates global contextual information with multiple local features to predict the optimal entity labels. Extensive experiments conducted on three datasets collected from English social media platforms (WNUT2016, WNUT2017, OntoNotes5.0\_English) demonstrate the advantageous performance of our proposed method, and ablation experiments further confirm the method's viability and effectiveness.  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2045-2322  
入藏号:  
WOS:001356224700032  
PubMed ID:  
39548167  
Indexed Date:  
2024-11-24  
Institution Address:  
Informat Engn Univ, Inst Geospatial Informat, Zhengzhou 450001, Peoples R China  
Institution:

### 99. Learn from orientation prior for radiograph super-resolution:

摘要: Background and objective: High -resolution radiographic images play a pivotal role in the early diagnosis and treatment of skeletal muscle -related diseases. It is promising to enhance image quality by introducing singleimage super -resolution (SISR) model into the radiology image field. However, the conventional image pipeline, which can learn a mixed mapping between SR and denoising from the color space and inter -pixel patterns, poses a particular challenge for radiographic images with limited pattern features. To address this issue, this paper introduces a novel approach: Orientation Operator Transformer - O2former. Methods: We incorporate an orientation operator in the encoder to enhance sensitivity to denoising mapping and to integrate orientation prior. Furthermore, we propose a multi -scale feature fusion strategy to amalgamate features captured by different receptive fields with the directional prior, thereby providing a more effective latent representation for the decoder. Based on these innovative components, we propose a transformer -based SISR model, i.e., O2former, specifically designed for radiographic images. Results: The experimental results demonstrate that our method achieves the best or second-best performance in the objective metrics compared with the competitors at x4 upsampling factor. For qualitative, more objective details are observed to be recovered. Conclusions: In this study, we propose a novel framework called O2former for radiological image super -resolution tasks, which improves the reconstruction model's performance by introducing an orientation operator and multi -scale feature fusion strategy. Our approach is promising to further promote the radiographic image enhancement field.  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0169-2607  
eISSN:  
1872-7565  
入藏号:  
WOS:001166228700001  
PubMed ID:  
38237449  
Indexed Date:  
2024-03-04  
Institution Address:  
Tohoku Univ, Grad Sch Engn, Dept Commun Engneering, Sendai 9808579, Japan  
Institution Address:  
Harvard Med Sch, Gordon Ctr Med Imaging, Boston, MA 02114 USA  
Institution Address:  
Tohoku Univ, Grad Sch Med, Dept Surg, Sendai 9808575, Japan  
Institution:  
Institution:  
Institution:

### 100. Multi-scale Feature Fusion and Transformer Network for urban green space

摘要: Accurate extraction of urban green space is critical for preserving urban ecological balance and enhancing urban life quality. However, due to the complex urban green space morphology (e.g., different sizes and shapes), it is still challenging to extract green space effectively from high-resolution image. To address this issue, we proposed a novel hybrid method, Multi-scale Feature Fusion and Transformer Network (MFFTNet), as a new deep learning approach for extracting urban green space from high-resolution (GF-2) image. Our method was characterized by two aspects: (1) a multi-scale feature fusion module and transformer network that enhanced the recovery of green space edge information and (2) vegetation feature (NDVI) that highlighted vegetation information and enhanced vegetation boundaries identification. The GF-2 image was utilized to build two urban green space labeled datasets, namely Greenfield and Greenfield2. We compared the proposed MFFTNet with the existing popular deep learning models (like PSPNet, DensASPP, etc.) to evaluate the effectiveness of MFFTNet by the Mean Intersection Over Union (MIOU) benchmark on Greenfield, Greenfield2, and a public dataset (WHDLD). Experiments on Greenfield2 showed that MFFTNet can achieve a high MIOU (86.50%), which outperformed deep learning networks like PSPNet and DensASPP by 0.86% and 3.28%, respectively. Meanwhile, the MIOU of MFFTNet incorporating vegetation feature (NDVI) was further achieved to 86.76% on Greenfield2. Our experimental results demonstrate that the proposed MFFTNet with vegetation feature (NDVI) outperforms the state-ofthe-art methods in urban green space segmentation.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
1  
Web of Science 核心合集中的 "被引频次":  
25  
被引频次合计:  
28  
ISSN:  
1569-8432  
eISSN:  
1872-826X  
入藏号:  
WOS:001098939300001  
Indexed Date:  
2023-11-27  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Nanjing 210044, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Geog Sci & Nat Resources Res, State Key Lab Resources & Environm Informat Syst, Beijing 100101, Peoples R China  
Institution Address:  
China Ctr Resources Satellite Data & Applicat, Beijing 100094, Peoples R China  
Institution Address:  
Minist Educ, Key Lab Poyang Lake Wetland & Watershed Res, Nanchang 330022, Peoples R China  
Institution Address:  
Jiangxi Normal Univ, Sch Geog & Environm, Nanchang 330022, Peoples R China  
Institution Address:  
UK Ctr Ecol & Hydrol, Lib Ave, Lancaster LA1 4AP, England  
Institution Address:  
Univ Bristol, Sch Geog Sci, Bristol BS8 1SS, England  
Institution:  
Institution:  
Institution:  
China Ctr Resources Satellite Data & Applicat  
Institution:  
Minist Educ  
Institution:  
Institution:  
Institution:

### 101. CXR-Seg: A Novel Deep Learning Network for Lung Segmentation from Chest

摘要: Over the past decade, deep learning techniques, particularly neural networks, have become essential in medical imaging for tasks like image detection, classification, and segmentation. These methods have greatly enhanced diagnostic accuracy, enabling quicker identification and more effective treatments. In chest X-ray analysis, however, challenges remain in accurately segmenting and classifying organs such as the lungs, heart, diaphragm, sternum, and clavicles, as well as detecting abnormalities in the thoracic cavity. Despite progress, these issues highlight the need for improved approaches to overcome segmentation difficulties and enhance diagnostic reliability. In this context, we propose a novel architecture named CXR-Seg, tailored for semantic segmentation of lungs from chest X-ray images. The proposed network mainly consists of four components, including a pre-trained EfficientNet as an encoder to extract feature encodings, a spatial enhancement module embedded in the skip connection to promote the adjacent feature fusion, a transformer attention module at the bottleneck layer, and a multi-scale feature fusion block at the decoder. The performance of the proposed CRX-Seg was evaluated on four publicly available datasets (MC, Darwin, and Shenzhen for chest X-rays, and TCIA for brain flair segmentation from MRI images). The proposed method achieved a Jaccard index, Dice coefficient, accuracy, sensitivity, and specificity of 95.63%, 97.76%, 98.77%, 98.00%, and 99.05%on MC; 91.66%, 95.62%, 96.35%, 95.53%, and 96.94% on V7 Darwin COVID-19; and 92.97%, 96.32%, 96.69%, 96.01%, and 97.40% on the Shenzhen Tuberculosis CXR Dataset, respectively. Conclusively, the proposed network offers improved performance in comparison with state-of-the-art methods, and better generalization for the semantic segmentation of lungs from chest X-ray images.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2306-5354  
入藏号:  
WOS:001431642800001  
PubMed ID:  
40001687  
Indexed Date:  
2025-03-03  
Institution Address:  
Texas A&M Univ, Elect & Comp Engn Program, Doha 23874, Qatar  
Institution Address:  
COMSATS Univ Islamabad, Dept Elect & Comp Engn, Abbottabad Campus, Abbottabad 22060, Pakistan  
Institution Address:  
Texas A&M Univ, Dept Elect & Comp Engn, College Stn, TX 77840 USA  
Institution:  
Institution:  
Institution:

### 102. Decoupled Cross-Modal Transformer for Referring Video Object

摘要: Referring video object segmentation (R-VOS) is a fundamental vision-language task which aims to segment the target referred by language expression in all video frames. Existing query-based R-VOS methods have conducted in-depth exploration of the interaction and alignment between visual and linguistic features but fail to transfer the information of the two modalities to the query vector with balanced intensities. Furthermore, most of the traditional approaches suffer from severe information loss in the process of multi-scale feature fusion, resulting in inaccurate segmentation. In this paper, we propose DCT, an end-to-end decoupled cross-modal transformer for referring video object segmentation, to better utilize multi-modal and multi-scale information. Specifically, we first design a Language-Guided Visual Enhancement Module (LGVE) to transmit discriminative linguistic information to visual features of all levels, performing an initial filtering of irrelevant background regions. Then, we propose a decoupled transformer decoder, using a set of object queries to gather entity-related information from both visual and linguistic features independently, mitigating the attention bias caused by feature size differences. Finally, the Cross-layer Feature Pyramid Network (CFPN) is introduced to preserve more visual details by establishing direct cross-layer communication. Extensive experiments have been carried out on A2D-Sentences, JHMDB-Sentences and Ref-Youtube-VOS. The results show that DCT achieves competitive segmentation accuracy compared with the state-of-the-art methods.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
1424-8220  
入藏号:  
WOS:001305101300001  
PubMed ID:  
39205068  
Indexed Date:  
2024-09-09  
Institution Address:  
Peoples Publ Secur Univ China, Sch Informat & Cyber Secur, Beijing 100038, Peoples R China  
Institution Address:  
Minist Publ Secur, Key Lab Secur Prevent Technol & Risk Assessment, Beijing 100038, Peoples R China  
Institution:  
Institution:

### 103. CBS-YOLOv5: fault detection algorithm of electrolyzer plate with

摘要: In the process of copper electrorefining, accurate detection of electrode plate faults is extremely challenging due to the low resolution of captured infrared images, significant noise interference, and dense electrode plate arrangements. To address these challenges, this paper proposes an improved YOLOv5-based electrode plate fault detection algorithm called CBS-YOLOv5. This algorithm introduces several innovations over the original YOLOv5, including: the incorporation of coordinate attention to enhance the ability of the feature extraction network to separate target information from noise; the construction of a small object detection module to improve the detection of dense small objects by increasing the resolution of the feature map; the replacement of the traditional path aggregation network with a Bi-directional Feature Pyramid Network (BiFPN) for more flexible multi-scale feature fusion; and the integration of the swin transformer to optimize the cross-stage partial bottleneck structure, significantly enhancing the model's ability to detect densely packed small objects. Experimental results show that the proposed CBS-YOLOv5 model achieves an accuracy of 88.1%, which is an improvement of 5.7% over the base model. Furthermore, this algorithm demonstrates exceptional detection capabilities for dense small objects in low-resolution infrared images while maintaining real-time detection speed, making it suitable for various complex industrial scenarios, including fault detection in non-ferrous metal electrolysis processes. CBS-YOLOv5 not only improves detection accuracy and robustness but also has broad application prospects, offering a new solution for intelligent manufacturing and industrial inspection.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001335507000001  
Indexed Date:  
2024-10-25  
Institution Address:  
Cent South Univ, Sch Automat, Changsha, Peoples R China  
Institution Address:  
Guangxi Univ, Sch Comp & Elect Informat, Nanning 530004, Peoples R China  
Institution:  
Institution:

### 104. Stage-Aware Interaction Network for Point Cloud Completion

摘要: Point cloud completion aims to restore full shapes of objects from partial scans, and a typical network pipeline is AutoEncoder, which has coarse-to-fine refinement modules. Although existing approaches using this kind of architecture achieve promising results, they usually neglect the usage of shallow geometry features in partial inputs and the fusion of multi-stage features in the upsampling process, which prevents network performances from further improving. Therefore, in this paper, we propose a new method with dense interactions between different encoding and decoding steps. First, we introduce the Decoupled Multi-head Transformer (DMT), which implements and integrates semantic prediction and resolution upsampling in a unified network module, which serves as a primary ingredient in our pipeline. Second, we propose an Encoding-aware Coarse Decoder (ECD) that compactly makes the top-down shape-decoding process interact with the bottom-up feature-encoding process to utilize both shallow and deep features of partial inputs for coarse point cloud generation. Third, we design a Stage-aware Refinement Group (SRG), which comprehensively understands local semantics from densely connected features across different decoding stages and gradually upsamples point clouds based on them. In general, the key contributions of our method are the DMT for joint semantic-resolution generation, the ECD for multi-scale feature fusion-based shape decoding, and the SRG for stage-aware shape refinement. Evaluations on two synthetic and three real-world datasets illustrate that our method achieves competitive performances compared with existing approaches.  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2079-9292  
入藏号:  
WOS:001305909200001  
Indexed Date:  
2024-09-15  
Institution Address:  
Shanghai Jiao Tong Univ, Sch Mech Engn, Shanghai 200240, Peoples R China  
Institution:

### 105. YOLO-T: Multi-Target Detection Algorithm for Transmission Lines

摘要: During UAV inspections of transmission lines, inspectors often encounter long distance and obstructed targets. However, existing detection algorithms tend to be less accurate when trying to detect these targets. Existing algorithms perform inadequately in handling long-distance and occluded targets, lacking effective detection capabilities for small objects and complex backgrounds. Therefore, we propose an improved YOLOv8-based YOLO-T algorithm for detecting multiple targets on transmission lines, optimized using transfer learning. Firstly, the model is lightweight while ensuring detection accuracy by replacing the original convolution block in the C2f module of the neck network with Ghost convolution. Secondly, to improve the target detection ability of the model, the C2f module in the backbone network is replaced with the Contextual Transformer module. Then, the feature extraction of the model is improved by integrating the Attention module and the residual edge on the SPPF (Spatial Pyramid Pooling-Fast). Finally, we introduce a new shallow feature layer to enable multi-scale feature fusion, optimizing the model detection accuracy for small and obscured objects. Parameters and GFLOPs are conserved by using the Add operation instead of the Concat operation. The experiment reveals that the enhanced algorithm achieves a mean detection accuracy of 97.19% on the transmission line dataset, which is 2.03% higher than the baseline YOLOv8 algorithm. It can also effectively detect small and occluded targets at long distances with a high FPS (98.91 frames/s).  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
2158-107X  
eISSN:  
2156-5570  
入藏号:  
WOS:001315627600108  
Indexed Date:  
2024-09-23  
Institution Address:  
Shanghai Dianji Univ, Sch Mech Engn, Shanghai 201306, Peoples R China  
Institution:

### 106. MRF-Net: An Infrared Remote Sensing Image Thin Cloud Removal Method With

摘要: The usability of infrared remote sensing data is often compromised by thin cloud cover. To address this problem, we proposed the multiscale residual fusion network (MRF-Net) to remove thin cloud from infrared remote sensing imagery. Initially, we developed a thin cloud simulation method utilizing Perlin noise and affine transformation to generate high-fidelity thin cloud representations. Subsequently, to accurately discern and eliminate thin cloud from infrared images, we proposed MRF-Net. This model incorporates a multiscale feature fusion module (MSFFM) for extracting shallow features, a residual dense network module (RDNM) for in-depth feature extraction, a residual Swin transformer module (RSTM) for capturing global features, and attention mechanisms to selectively enhance target information. The Swin transformer, a hierarchical Transformer whose representation is computed with shifted windows, is employed to improve the efficiency of global feature extraction. Finally, we devised a combined loss function that accounts for both intrablock and interblock constraints to ensure de-clouding consistency across different image blocks. The intrablock constraint focuses on removing thin cloud within each image block, while the interblock constraint is designed to enhance the consistency of cloud removal between blocks. We have assembled a dataset comprising both simulated and real data to validate the efficiency of our proposed method. Experimental results have shown that our method effectively eliminates thin cloud and surpasses existing state-of-the-art methods. The source codes are available at https://github.com/CastleChen339/MRF-Net.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001399064000020  
Indexed Date:  
2025-01-29  
Institution Address:  
Beijing Inst Technol, Sch Mechtron Engn, Beijing 100081, Peoples R China  
Institution Address:  
Beijing Inst Technol, Sch Informat & Elect, Beijing 100081, Peoples R China  
Institution:  
Institution:

### 107. OD-YOLO: Robust Small Object Detection Model in Remote Sensing Image

摘要: As remote sensing technology has advanced, the use of satellites and similar technologies has become increasingly prevalent in daily life. Now, it plays a crucial role in hydrology, agriculture, and geography. Nevertheless, because of the distinct qualities of remote sensing, including expansive scenes and small, densely packed targets, there are many challenges in detecting remote sensing objects. Those challenges lead to insufficient accuracy in remote sensing object detection. Consequently, developing a new model is essential to enhance the identification capabilities for objects in remote sensing imagery. To solve these constraints, we have designed the OD-YOLO approach that uses multi-scale feature fusion to improve the performance of the YOLOv8n model in small target detection. Firstly, traditional convolutions have poor recognition capabilities for certain geometric shapes. Therefore, in this paper, we introduce the Detection Refinement Module (DRmodule) into the backbone architecture. This module utilizes Deformable Convolutional Networks and the Hybrid Attention Transformer to strengthen the model's capability for feature extraction from geometric shapes and blurred objects effectively. Meanwhile, based on the Feature Pyramid Network of YOLO, at the head of the model framework, this paper enhances the detection capability by introducing a Dynamic Head to strengthen the fusion of different scales features in the feature pyramid. Additionally, to address the issue of detecting small objects in remote sensing images, this paper specifically designs the OIoU loss function to finely describe the difference between the detection box and the true box, further enhancing model performance. Experiments on the VisDrone dataset show that OD-YOLO surpasses the compared models by at least 5.2% in mAP50 and 4.4% in mAP75, and experiments on the Foggy Cityscapes dataset demonstrated that OD-YOLO improved mAP by 6.5%, demonstrating outstanding results in tasks related to remote sensing images and adverse weather object detection. This work not only advances the research in remote sensing image analysis, but also provides effective technical support for the practical deployment of future remote sensing applications.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
3  
eISSN:  
1424-8220  
入藏号:  
WOS:001245745000001  
PubMed ID:  
38894387  
Indexed Date:  
2024-06-19  
Institution Address:  
Zhejiang Sci Tech Univ, Sch Comp Sci & Technol, Hangzhou 310018, Peoples R China  
Institution Address:  
Zhejiang Sci Tech Univ, KeYi Coll, Shaoxing 312369, Peoples R China  
Institution:  
Institution:

### 108. TGSYOLO: Template-Guidance Siamese Network for SMT Welding Defect

摘要: Surface-mounted technology (SMT) welding defect detection plays a key role in the printed circuit board assembly (PCBA) production process, which affects the use of electronic products and cost. Previous works tend to realize defect detection with only defect samples and they assume that there are sufficient defect samples. However, defect samples are usually difficult to collect in real-life scenarios while enough template samples can be easily obtained. In addition, most existing works carry out defect detection based on benchmarks with simple backgrounds of PCBA, which is not suitable for PCBA with complex structures in modern electronic product manufacturing. To address the above issues, we propose a template-guidance Siamese network based on YOLO for SMT welding defect detection (TGSYOLO), which is deployed on a real SMT automatic optical inspection (AOI) system. First, the two-stream structure is introduced to extract deep features in defect images and template images, in which template features serve as guidance knowledge. Then, a template fusion Transformer (TFT) is proposed to model global features between detect and template features in the low-level stage, which could acquire long-range correlations to force the network to focus on potential defect regions. Next, to avoid the disappearance of tiny defect features during deep feature fusion, a multiscale attention feature pyramid network (MAFPN) is proposed to directly fuse defect semantic information from low-level features, which retains detailed expressions of defects through skip connection and obtains compact fusion features. Furthermore, we collect limited welding defect samples based on more complex PCBA backgrounds than previous works through a real SMT AOI system. Experiments on the limited dataset show that TGSYOLO could reach 0.985 of mAP@0.5, 0.885 of mAP@0.75, and 0.984 of F1, which is 0.008, 0.054, and 0.025 higher than other SOTA methods. Also, generalization experiments on the public DeepPCB show that TGSYOLO could still reach the best with 0.991 of mAP@0.5 and 0.89 of mAP@0.75, which proves that TGSYOLO has good generalization performance.  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
2156-3950  
eISSN:  
2156-3985  
入藏号:  
WOS:001389745800012  
Indexed Date:  
2025-01-11  
Institution Address:  
Univ Sci & Technol China, Dept Automat, Hefei 230026, Peoples R China  
Institution Address:  
Minist Educ, Key Lab Syst Control & Informat Proc, Shanghai 200240, Peoples R China  
Institution Address:  
Yangtze River Delta Hart Robot Ind Technol Res Ins, Wuhu 241000, Peoples R China  
Institution Address:  
Anhui Univ, AHU IAI AI Joint Lab, Hefei 230601, Peoples R China  
Institution Address:  
Hefei Comprehens Natl Sci Ctr, Inst Artificial Intelligence, Hefei 230088, Peoples R China  
Institution Address:  
Chinese Acad Sci, Key Lab Technol GeoSpatial Informat Proc & Applica, Beijing 100192, Peoples R China  
Institution:  
Institution:  
Minist Educ  
Institution:  
Yangtze River Delta Hart Robot Ind Technol Res Ins  
Institution:  
Institution:  
Hefei Comprehens Natl Sci Ctr  
Institution:

### 109. Optimizing feature extraction and fusion for high-resolution defect

摘要: In this paper, we propose a novel architecture for defect detection in electroluminescent images of polycrystalline silicon solar cells, addressing the challenges posed by subtle and dispersed defects. Our model, based on a modified Swin Transformer, incorporates key innovations that enhance feature extraction and fusion. We replace the conventional self-attention mechanism with a novel group self-attention mechanism, increasing the mAP50:5:95 score from 50.12 % to 52.98 % while reducing inference time from 74 ms to 62 ms. We also introduce a spatial displacement with shift convolution module, replacing the traditional Multi-Layer Perceptron, which further enhances the model's receptive field and improves precision and recall. Additionally, our fast multi-scale feature fusion mechanism effectively combines high-resolution details with high-level semantic features from different network layers, optimizing defect detection accuracy. Experimental results on the PVEL-AD dataset demonstrate that our model achieves the highest mAP50 score of 83.11 % and an F1-Score of 84.33 %, surpassing state-of-the-art models while maintaining a competitive inference time of 66.3 ms. These findings highlight the effectiveness of our innovations in improving defect detection accuracy and computational efficiency, making our model a robust solution for quality assurance in solar cell manufacturing.  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2667-3053  
入藏号:  
WOS:001330931200001  
Indexed Date:  
2024-10-17  
Institution Address:  
Ind Univ Ho Chi Minh City, Fac Elect Engn Technol, Ho Chi Minh City 700000, Vietnam  
Institution:

### 110. A Semantically Guided Deep Supervised Hashing Model for Multi-Label

摘要: With the rapid growth of remote sensing data, efficiently managing and retrieving large-scale remote sensing images has become a significant challenge. Specifically, for multi-label image retrieval, single-scale feature extraction methods often fail to capture the rich and complex information inherent in these images. Additionally, the sheer volume of data creates challenges in retrieval efficiency. Furthermore, leveraging semantic information for more accurate retrieval remains an open issue. In this paper, we propose a multi-label remote sensing image retrieval method based on an improved Swin Transformer, called Semantically Guided Deep Supervised Hashing (SGDSH). The method aims to enhance feature extraction capabilities and improve retrieval precision. By utilizing multi-scale information through an end-to-end learning approach with a multi-scale feature fusion module, SGDSH effectively integrates both shallow and deep features. A classification layer is introduced to assist in training the hash codes, incorporating RS image category information to improve retrieval accuracy. The model is optimized for multi-label retrieval through a novel loss function that combines classification loss, pairwise similarity loss, and hash code quantization loss. Experimental results on three publicly available remote sensing datasets, with varying sizes and label distributions, demonstrate that SGDSH outperforms state-of-the-art multi-label hashing methods in terms of average accuracy and weighted average precision. Moreover, SGDSH returns more relevant images with higher label similarity to query images. These findings confirm the effectiveness of SGDSH for large-scale remote sensing image retrieval tasks and provide new insights for future research on multi-label remote sensing image retrieval.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001442484700001  
Indexed Date:  
2025-03-18  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100094, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Sch Elect Elect & Commun Engn, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 111. Efficient X-ray Security Images for Dangerous Goods Detection Based on

摘要: In response to the problems of complex background, multi-scale dangerous goods and severe stacking in X-ray security images, this paper proposes a high-accuracy dangerous goods detection algorithm for X-ray security images based on the improvement of YOLOv7. Firstly, by combining the coordinate attention mechanism, the downsampling structure of the backbone network is improved to enhance the model's target feature localization ability. Secondly, a weighted bidirectional feature pyramid network is used as the feature fusion structure to achieve multi-scale feature weighted fusion and further simplify the network. Then, combined with dynamic snake convolution, a downsampling structure was designed to facilitate the extraction of features at different scales, providing richer feature representations. Finally, drawing inspiration from the idea of group convolution and combining it with Conv2Former, a feature extraction module called a multi-convolution transformer (MCT) was designed to enhance the network's feature extraction ability by combining multi-scale information. The improved YOLOv7 in this article was tested on the public datasets SIXRay, CLCXray, and PIDray. The average detection accuracy (mAP) of the improved model was 96.3%, 79.3%, and 84.7%, respectively, which was 4.7%, 2.7%, and 3.1% higher than YOLOv7. This proves the effectiveness and universality of the method proposed in this article. Compared to the current mainstream X-ray image dangerous goods detection models, this model effectively reduces the false detection rate of dangerous goods in X-ray security inspection images and has achieved significant improvement in the detection of small and multi-scale targets, achieving higher accuracy in dangerous goods detection.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2079-9292  
入藏号:  
WOS:001210581900001  
Indexed Date:  
2024-05-10  
Institution Address:  
Univ Elect Sci & Technol China, Zhongshan Inst, Coll Electron & Informat, Zhongshan 528402, Peoples R China  
Institution Address:  
Harbin Univ Sci & Technol, Heilongjiang Prov Key Lab Laser Spect Technol & Ap, Harbin 150080, Peoples R China  
Institution:  
Institution:

### 112. MF-Net: a multimodal fusion network for emotion recognition based on

摘要: Currently, research on emotion recognition has shown that multi-modal data fusion has advantages in improving the accuracy and robustness of human emotion recognition, outperforming single-modal methods. Despite the promising results of existing methods, significant challenges remain in effectively fusing data from multiple modalities to achieve superior performance. Firstly, existing works tend to focus on generating a joint representation by fusing multi-modal data, with fewer methods considering the specific characteristics of each modality. Secondly, most methods fail to fully capture the intricate correlations among multiple modalities, often resorting to simplistic combinations of latent features. To address these challenges, we propose a novel fusion network for multi-modal emotion recognition. This network enhances the efficacy of multi-modal fusion while preserving the distinct characteristics of each modality. Specifically, a dual-stream multi-scale feature encoding (MFE) is designed to extract emotional information from both electroencephalogram (EEG) and peripheral physiological signals (PPS) temporal slices. Subsequently, a cross-modal global-local feature fusion module (CGFFM) is proposed to integrate global and local information from multi-modal data and then assign different importance to each modality, which makes the fusion data tend to the more important modalities. Meanwhile, the transformer module is employed to further learn the modality-specific information. Moreover, we introduce the adaptive collaboration block (ACB), which optimally leverages both modality-specific and cross-modality relations for enhanced integration and feature representation. Following extensive experiments on the DEAP and DREAMER multimodal datasets, our model achieves state-of-the-art performance.  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1863-1703  
eISSN:  
1863-1711  
入藏号:  
WOS:001370437500007  
Indexed Date:  
2024-12-11  
Institution Address:  
Hangzhou Dianzi Univ, Sch Automat, Hangzhou 310000, Peoples R China  
Institution Address:  
Hangzhou City Univ, Sch Med, Hangzhou 310015, Peoples R China  
Institution Address:  
Hangzhou Dianzi Univ, Sch Comp Sci, Hangzhou 310000, Peoples R China  
Institution Address:  
Hangzhou City Univ, Key Lab Brain Machine Collaborat Intelligence Zhej, Hangzhou 310015, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 113. EDFIDepth: enriched multi-path vision transformer feature interaction

摘要: Monocular depth estimation (MDE) aims to predict pixel-level dense depth maps from a single RGB image. Some recent approaches mainly rely on encoder-decoder architectures to capture and process multi-scale features. However, they usually exploit heavier network at the expense of computational costs to obtain high-quality depth maps. In this paper, we propose a novel enriched multi-path vision transformer feature interaction network with an encoder-decoder architecture, denoted as EDFIDepth , which seeks a balance between computational costs and performance rather than pursuing the highest accuracy or extremely lightweight models. Specifically, an encoder called MPViT-D, incorporating multi-path vision transformer and a deep convolution module, is introduced to extract diverse features with both fine and coarse details at the same feature level with fewer parameters. Subsequently, we propose a lightweight decoder comprising two effective modules to establish multi-scale feature interaction: an encoder-decoder cross-feature matching (ED-CFM) module and a channel-level feature fusion (CLFF) module. The ED-CFM module is to establish connections between encoder-decoder features through a dual-path structure, where a cross-attention mechanism is deployed to enhance the relevance of multi-scale complementary depth information. Meanwhile, the CLFF module utilizes a channel attention mechanism to further fuse crucial depth information within the channels, thereby improving the accuracy of depth estimation. Extensive experiments on the indoor dataset NYUv2 and the outdoor dataset KITTI demonstrate that our method can achieve comparable state-of-the-art (SOTA) results while significantly reducing the number of trainable parameters. Our codes and approach are available at https://github.com/Zhangmg123/EDFIDEpth.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0920-8542  
eISSN:  
1573-0484  
入藏号:  
WOS:001242163900001  
Indexed Date:  
2024-06-14  
Institution Address:  
Anhui Univ Sci & Technol, Coll Comp Sci & Engn, Huainan 232001, Peoples R China  
Institution Address:  
Anhui Univ Sci & Technol, Huainan Peoples Hosp 1, Affiliated Hosp 1, Huainan, Anhui, Peoples R China  
Institution Address:  
Anhui Purvar Bigdata Technol Co Ltd, Huainan 232001, Peoples R China  
Institution Address:  
Anhui Univ Sci & Technol, Coll Elect & Informat Engn, Huainan, Anhui, Peoples R China  
Institution Address:  
Providence Univ, Dept Comp Sci & Informat Engn, Taichung, Taiwan  
Institution Address:  
Anhui Univ, Sch Elect & Informat Engn, Hefei, Anhui, Peoples R China  
Institution Address:  
Anhui Univ Sci & Technol, Inst Environm friendly Mat & Occupat Hlth, Wuhu, Anhui, Peoples R China  
Institution:  
Institution:  
Institution:  
Anhui Purvar Bigdata Technol Co Ltd  
Institution:  
Institution:  
Institution:  
Institution:

### 114. An Unsupervised Fundus Image Enhancement Method with Multi-Scale

摘要: Color fundus images are now widely used in computer-aided analysis systems for ophthalmic diseases. However, fundus imaging can be affected by human, environmental, and equipment factors, which may result in low-quality images. Such quality fundus images will interfere with computer-aided diagnosis. Existing methods for enhancing low-quality fundus images focus more on the overall visualization of the image rather than capturing pathological and structural features at the finer scales of the fundus image sufficiently. In this paper, we design an unsupervised method that integrates a multi-scale feature fusion transformer and an unreferenced loss function. Due to the loss of microscale features caused by unpaired training, we construct the Global Feature Extraction Module (GFEM), a combination of convolution blocks and residual Swin Transformer modules, to achieve the extraction of feature information at different levels while reducing computational costs. To improve the blurring of image details caused by deep unsupervised networks, we define unreferenced loss functions that improve the model's ability to suppress edge sharpness degradation. In addition, uneven light distribution can also affect image quality, so we use an a priori luminance-based attention mechanism to improve low-quality image illumination unevenness. On the public dataset, we achieve an improvement of 0.88 dB in PSNR and 0.024 in SSIM compared to the state-of-the-art methods. Experiment results show that our method outperforms other deep learning methods in terms of vascular continuity and preservation of fine pathological features. Such a framework may have potential medical applications.  
Web of Science 核心合集中的 "被引频次":  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2079-9292  
入藏号:  
WOS:001028170600001  
Indexed Date:  
2023-07-28  
Institution Address:  
Wuhan Text Univ, Sch Elect & Elect Engn, Wuhan 430077, Peoples R China  
Institution Address:  
Wuhan Univ, Sch Comp Sci, Wuhan 430072, Peoples R China  
Institution Address:  
Tarim Univ, Sch Informat Engn, Alaer 843300, Peoples R China  
Institution:  
Institution:  
Institution:

### 115. Context-aware semantic segmentation network for tunnel face feature

摘要: The automated interpretation of tunnel face geological information is significant to the construction decisionmaking of rock mass engineering. An intelligent recognition algorithm, named the Transformer and Convolution neural networks Semantic segmentation Network (TCSeNet), is introduced to overcome the low interpretation accuracy caused by certain limitations of existing automated interpretation methods. Firstly, a hybrid encoder is constructed to extract global and local context information using Transformer and Convolution Neural Networks, respectively. Furthermore, a context feature adaptive selection module is designed to cross-fusion the information. Secondly, a feature pyramid-like decoder is constructed for multi-scale object problems. Furthermore, a multi-scale feature refinement module is designed to improve the ability of multi-scale feature expression. A tunnel face dataset is constructed for training and testing. The research shows that the mIoU of TCSeNet reaches 92.57%. This indicates that the TCSeNet is suitable for automatic interpretation of tunnel face geological information.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0926-5805  
eISSN:  
1872-7891  
入藏号:  
WOS:001256286500001  
Indexed Date:  
2024-07-04  
Institution Address:  
Xian Univ Architecture & Technol, Coll Informat & Control Engn, Xian 710055, Shaanxi, Peoples R China  
Institution Address:  
Shaanxi Key Lab Geotech & Underground Space Engn, Xian 710055, Shaanxi, Peoples R China  
Institution Address:  
Xian Univ Architecture & Technol, Coll Civil Engn, Xian 710055, Shaanxi, Peoples R China  
Institution:  
Institution:  
Shaanxi Key Lab Geotech & Underground Space Engn  
Institution:

### 116. MSS-Former: Multiscale Skeletal Transformer for Intelligent Fall Risk

摘要: Fall, a leading cause of accidental death and injury in older adults aged 65 and above, has become a rapidly growing health concern in aging populations worldwide. Data-driven methods integrating depth imaging technology have received growing attention in automated fall risk assessment owing to their noninvasiveness and less dependence on healthcare professionals. However, most existing depth image data-based models neglect the inherent physiological and potential functional connections and lack sufficient real-world data validation. To fill the research gap, we developed a novel approach named multiscale skeletal transformer (MSS-Former), leveraging depth image technology and deep-learning models for effective fall risk prediction. Our contributions mainly consist of four parts. First, we introduced a multimodel output feature fusion transformer in fall risk prediction, enabling output merging and weighting from multiple model streams dynamically. Second, we developed an innovative scheme to construct interjoint skeletal topology, systematically focusing on joints' intrinsic physiological and potential functional connections. Third, we constructed a ResNet-FPN, greatly enhancing multiscale feature extraction capabilities. Fourth, we conducted a field study in a local hospital and performed a comprehensive validation of our developed approach. The comparison results show that our approach achieved outstanding predictive performance, surpassing state-of-the-art methods on the real-world data set, with accuracy, precision, recall, and F1 scores of 97.84%, 97.33%, 96.97%, and 96.92%, respectively. In practice, the proposed approach would be of great value in the timely identification for individuals at high fall risk and facilitate decision making to take appropriate interventions.  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
2327-4662  
入藏号:  
WOS:001330865200043  
Indexed Date:  
2024-10-18  
Institution Address:  
Sun Yat sen Univ, Intelligent Sensing & Proact Hlth Res Ctr, Sch Publ Hlth Shenzhen, Shenzhen 518000, Peoples R China  
Institution Address:  
Shenzhen Technol Univ, Coll Big Data & Internet, Shenzhen 518118, Peoples R China  
Institution Address:  
Univ Hong Kong, Shenzhen Hosp, Dept Physiotherapy, Shenzhen 518000, Peoples R China  
Institution Address:  
Virginia Polytech Inst & State Univ, Grado Dept Ind & Syst Engn, Blacksburg, VA 24061 USA  
Institution:  
Institution:  
Institution:  
Institution:

### 117. When Transformer Meets Robotic Grasping: Exploits Context for Efficient

摘要: In this letter, we present a transformer-based architecture, namely TF-Grasp, for robotic grasp detection. The developed TF-Grasp framework has two elaborate designs making it well suitable for visual grasping tasks. The first key design is that we adopt the local window attention to capture local contextual information and detailed features of graspable objects. Then, we apply the cross window attention to model the long-term dependencies between distant pixels. Object knowledge, environmental configuration, and relationships between different visual entities are aggregated for subsequent grasp detection. The second key design is that we build a hierarchical encoder-decoder architecture with skip-connections, delivering shallow features from the encoder to decoder to enable a multi-scale feature fusion. Due to the powerful attention mechanism, TF-Grasp can simultaneously obtain the local information (i.e., the contours of objects), and model long-term connections such as the relationships between distinct visual concepts in clutter. Extensive computational experiments demonstrate that TF-Grasp achieves competitive results versus state-of-art grasping convolutional models and attains a higher accuracy of 97.99 and 94.6% on Cornell and Jacquard grasping datasets, respectively. Real-world experiments using a 7DoF Franka Emika Panda robot also demonstrate its capability of grasping unseen objects in a variety of scenarios.  
Web of Science 核心合集中的 "被引频次":  
57  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
6  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
62  
ISSN:  
2377-3766  
入藏号:  
WOS:000838409600007  
Indexed Date:  
2022-08-24  
Institution Address:  
Univ Sci & Technol China, Dept Automat, Hefei 230026, Peoples R China  
Institution:

### 118. HP-CRL: High-resolution preservation driven collaborative representation

摘要: The Infrared and Visible Image Fusion (IVIF) task aims to generate a fused image that capitalizes on the salient features of the infrared image and the textural details of the visible image. The disparity between infrared and visible modalities has long been recognized as a significant impediment within the realm of IVIF. To address the challenge of integrating modality -specific and modality -shared features in cross -modality feature fusion, we propose a novel high -resolution preservation driven collaborative representation learning method to effectively fuse infrared and visible images, known as HP-CRL. In our model, a high -resolution preservation module is proposed to progressively extract multi -scale feature representations. More specifically, to alleviate the loss of features caused by down -sampling, we draw inspiration from the back -projection technique to continuously complement multi -resolution features while simultaneously maintaining high -resolution representations. Also, a module combining both Vision Transformer (Vit) and convolutional attention is employed to enhance the semantic representation of source images. Aiming at the problem of information redundancy during feature extraction, we employ a multi -branch transmission module for collaborative representation learning across branches and full interaction between multi -scale features. Our experiments demonstrate the efficacy of HPCRL, surpassing other 15 state-of-the-art (SOTA) fusion methods. The results suggest the promise of our approach in achieving superior fusion quality and maintaining the salient characteristics of the source images.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0030-3992  
eISSN:  
1879-2545  
入藏号:  
WOS:001244595800001  
Indexed Date:  
2024-06-21  
Institution Address:  
Yunnan Univ, Sch Informat Sci & Engn, Kunming 650500, Peoples R China  
Institution Address:  
Southeast Univ, Sch Automat, Nanjing 211189, Peoples R China  
Institution Address:  
Southeast Univ, Sch Math, Nanjing 211189, Peoples R China  
Institution Address:  
Ahlia Univ, Manama, Bahrain  
Institution:  
Institution:  
Institution:  
Institution:

### 119. MSFFT: Multi-Scale Feature Fusion Transformer for cross platform vehicle

摘要: A vital component of Intelligent Transportation Systems (ITS) is vehicle re -identification, which allows vehicles to be identified across surveillance devices. Re -identification of vehicles is usually done using information collected from standalone surveillance devices such as fixed surveillance cameras (CCTVs) or aerial devices (UAVs). Re -identifying vehicles across standalone surveillance systems is challenging when there is a severe illumination change, a change of viewpoint, or an occlusion. Cross platform surveillance (CCTV+UAV) based vehicle re -identification is yet to be explored and can mitigate some of the challenges faced during reidentifying vehicles with standalone surveillance systems. This paper proposes a novel cross platform vehicle identification dataset called MCU-VReID using 42 CCTVs and a UAV. A novel re -identification method called Multi -Scale Feature Fusion Transformer (MSFFT) is proposed to re -identify vehicles observed across the cross platform surveillance systems. The network consists of inception layers with transformer networks that enable it to learn the vehicle's features at a variety of scales. The vehicles observed by two contrasting surveillance systems appear to be transformed representations of one another. Hence a two -stage training approach is facilitated for re -identifying vehicles observed across cross platform surveillance systems. The two -stage training approach aims to learn vehicle semantic transformations in the first stage using selfsupervised approaches. The knowledge gained at the first stage relating to vehicle semantic transformations is transferred at the second stage of training to perform re -identification. Extensive experiments using the method demonstrate that MSFFT significantly improves over state-of-the-art methods to perform cross platform vehicle re -identification.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:001220880100001  
Indexed Date:  
2024-05-18  
Institution Address:  
Manipal Inst Technol, Manipal Acad Higher Educ, Dept Informat & Commun Technol, Manipal 576104, Karnataka, India  
Institution Address:  
Manipal Acad Higher Educ, Manipal Inst Technol Bengaluru, Dept Elect & Commun Engn, Manipal 576104, Karnataka, India  
Institution Address:  
Manipal Inst Technol, Manipal Acad Higher Educ, Dept Data Sci & Comp Applicat, Manipal 576104, Karnataka, India  
Institution:  
Institution:  
Institution:

### 120. LiDAR-SLAM loop closure detection based on multi-scale point cloud

摘要: Loop closure detection is an important part of simultaneous localization and mapping (SLAM), used to detect and correct map and trajectory drift issues caused by accumulated errors. For the LiDAR-SLAM system, the sparsity and disorder of point clouds make position recognition based on 3D point clouds more challenging. In recent years, many deep learning-based closed-loop detection algorithms have been proposed. However, due to the sparsity of point clouds, current deep learning algorithms often convert point clouds into histograms or depth maps and then process them using deep learning algorithms, undoubtedly causing information loss. In this paper, we propose a closed-loop detection method based on multi-scale point cloud features transformer, which introduces multi-scale point cloud feature extraction and transformer global context modeling. We use voxel sparse convolution to obtain features of original point clouds at different resolutions and establish contextual relationships between features at different resolutions using the transformer network to achieve multi-scale feature fusion, and then obtain global descriptors. The obtained global descriptors can be used not only for closed-loop detection but also for front-end registration to address the challenges of point cloud processing in the SLAM system, especially in enhancing global modeling capabilities and reducing information loss. Our method directly processes point cloud data and integrates multi-scale point cloud feature information, which can better adapt to the characteristics of LiDAR-SLAM systems, improving the accuracy and robustness of localization and map construction, thus having broad application prospects in the field of measurement. We evaluated our method on multiple sequences of the KITTI and KITTI-360 datasets, each containing more than 5000 frames of point clouds, and extensively evaluated on a self-collected dataset of over 3.6 km. The experimental results show that our method achieves an accuracy of over 80% on multiple datasets and demonstrates superior performance in different environments.  
Web of Science 核心合集中的 "被引频次":  
6  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
6  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001131990300001  
Indexed Date:  
2024-01-11  
Institution Address:  
Jiangsu Univ, Inst Automot Res, Zhenjiang, Peoples R China  
Institution:

### 121. SS-DETR: a strong sensing DETR road obstacle detection model based on

摘要: As a key step in obstacle avoidance and path planning, obstacle detection via camera sensors is crucial for autonomous driving. The real traffic road environment is complex and variable, and the existing obstacle detection algorithms still have the problem of insufficient sensing ability. Therefore, this work suggests a camera sensors-based Strong Sensing DEtection TRansformer (SS-DETR) obstacle detection model for autonomous driving. Firstly, receptive-field attention ResNet is designed to improve feature analysis and extraction performance by considering the importance of receptive field spatial features and channels. Then, an intra-scale feature interaction module based on multiple information fusion attention is created to strengthen the representation of advanced feature maps. Furthermore, the cross-scale feature-fusion module is optimized to extract more detailed information from multi-scale feature maps. Finally, a localization loss function based on L1 and Powerful Intersection over Union v2 is implemented to further boost the detection performance. To verify the efficacy of the suggested model, the KITTI dataset containing camera sensors-based road obstacle images is adopted. The experimental results reveal that compared to real-time DETR, SS-DETR improves mean average precision (mAP)@50:95 and mAP@50 by 2.4% and 1.9%, respectively, and has a real-time inference speed of 33.7 frames per second. To further confirm the generalization ability of the approach, experiments are conducted on the camera sensors-based Cityscapes dataset. The results divulge that the suggested strategy can effectively raise the detection accuracy of obstacles, and offer a fresh perspective on obstacle identification.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0957-0233  
eISSN:  
1361-6501  
入藏号:  
WOS:001396542300001  
Indexed Date:  
2025-01-20  
Institution Address:  
Chongqing Univ, State Key Lab Mech Transmiss Adv Equipment, Chongqing 400044, Peoples R China  
Institution Address:  
Chongqing Univ, Coll Mech & Vehicle Engn, Chongqing 400044, Peoples R China  
Institution Address:  
Wuhan Inst Technol, Sch Mech & Elect Engn, Wuhan 430200, Peoples R China  
Institution Address:  
Chongqing Univ Technol, Coll Mech Engn, Chongqing 400044, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 122. RS-DETR: An Improved Remote Sensing Object Detection Model Based on

摘要: Object detection is a fundamental task in computer vision. Recently, deep-learning-based object detection has made significant progress. However, due to large variations in target scale, the predominance of small targets, and complex backgrounds in remote sensing imagery, remote sensing object detection still faces challenges, including low detection accuracy, poor real-time performance, high missed detection rates, and high false detection rates in practical applications. To enhance remote sensing target detection performance, this study proposes a new model, the remote sensing detection transformer (RS-DETR). First, we incorporate cascaded group attention (CGA) into the attention-driven feature interaction module. By capturing features at different levels, it enhances the interaction between features through cascading and improves computational efficiency. Additionally, we propose an enhanced bidirectional feature pyramid network (EBiFPN) to facilitate multi-scale feature fusion. By integrating features across multiple scales, it improves object detection accuracy and robustness. Finally, we propose a novel bounding box regression loss function, Focaler-GIoU, which makes the model focus more on difficult samples, improving detection performance for small and overlapping targets. Experimental results on the satellite imagery multi-vehicles dataset (SIMD) and the high-resolution remote sensing object detection (TGRS-HRRSD) dataset show that the improved algorithm achieved mean average precision (mAP) of 78.2% and 91.6% at an intersection over union threshold of 0.5, respectively, which is an improvement of 2.0% and 1.5% over the baseline model. This result demonstrates the effectiveness and robustness of our proposed method for remote sensing image object detection.  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
eISSN:  
2076-3417  
入藏号:  
WOS:001366881100001  
Indexed Date:  
2024-12-07  
Institution Address:  
China Univ Geosci, Sch Comp Sci, Wuhan 430074, Peoples R China  
Institution Address:  
China Univ Geosci, Informatizat Off, Wuhan 430074, Peoples R China  
Institution:  
Institution:

### 123. Computational spectral imaging reconstruction via a spatial-spectral

摘要: Compared with traditional hyperspectral imaging, computational spectral imaging (CSI) has the advantage of snapshot imaging with high spatial and temporal resolution, which has attracted considerable attention. The core challenge of CSI is to achieve computational imaging reconstruction from a single 2D measurement image to the corresponding 3D spatial-hyperspectral image (HSI). Existing reconstruction methods still face problems in exploring spatial-spectral cross correlation, leading to significant spatial-spectral distortion. Furthermore, due to neglect of multi-scale feature reconstruction, their reconstruction quality still needs to be improved. In this paper, to solve the above problems, we propose a spatial-spectral cross-attention-driven network (SSCA-DN). In SSCA, a proposed multi-scale feature aggregation (MFA) module and a spectral-wise transformer (SpeT) are used for multi-scale spatial feature reconstruction and long-range spectral feature reconstruction, respectively. Using spatial attention and spectral attention to interactively guide the reconstruction of the target HSI in spectral and spatial dimensions, the proposed SSCA models spatial-spectral cross correlation with considering multi-scale features. Using the SSCA as a basic module, a novel SSCA-DN network is constructed, in which a proposed supervised preliminary reconstruction subnetwork (SPRNet) learns the generalized prior, and a proposed unsupervised multiscale feature fusion and refinement subnetwork (UMFFRNet) learns the specific prior. The SSCA module ensures that the learned generalized and specific priors can capture the spatial-spectral cross correlation while considering multi-scale features. In addition, in UMFFRNet, driven by MFA and SSCA, a novel multi-scale fusion and refinement mechanism for multi-level adjacent features is proposed to effectively model the correlation between adjacent level features and the multi-scale spatial-spectral cross correlation, which further improves the reconstruction accuracy. Extensive experiments show that our method achieves state-of-the-art performance on both simulated and real datasets. (c) 2025 Optica Publishing Group. All rights, including for text and data mining (TDM), Artificial Intelligence (AI) training, and similar technologies, are reserved.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1084-7529  
eISSN:  
1520-8532  
入藏号:  
WOS:001416562800001  
Indexed Date:  
2025-02-20  
Institution Address:  
Beijing Inst Graph Commun, Sch Printing & Packaging, Beijing 102600, Peoples R China  
Institution Address:  
Beihang Univ, Sch Instrumentat & Optoelect Engn, Beijing 100191, Peoples R China  
Institution Address:  
Tongji Univ, Sch Phys Sci & Engn, Shanghai 200092, Peoples R China  
Institution:  
Institution:  
Institution:

### 124. Semantic-Aligned Matching for Enhanced DETR Convergence and Multi-Scale

摘要: The recently proposed DEtection TRansformer (DETR) has established a fully end-to-end paradigm for object detection. However, DETR suffers from slow training convergence, which hinders its applicability to various detection tasks. We observe that DETR's slow convergence is largely attributed to the difficulty in matching object queries to relevant regions due to the unaligned semantics between object queries and encoded image features. With this observation, we design Semantic-Aligned-Matching DETR++ (SAM-DETR++) to accelerate DETR's convergence and improve detection performance. The core of SAM-DETR++ is a plug-and-play module that projects object queries and encoded image features into the same feature embedding space, where each object query can be easily matched to relevant regions with similar semantics. Besides, SAM-DETR++ searches for multiple representative keypoints and exploits their features for semantic-aligned matching with enhanced representation capacity. Furthermore, SAM-DETR++ can effectively fuse multi-scale features in a coarse-to-fine manner on the basis of the designed semantic-aligned matching. Extensive experiments show that the proposed SAM-DETR++ achieves superior convergence speed and competitive detection accuracy. Additionally, as a plug-and-play method, SAM-DETR++ can complement existing DETR convergence solutions with even better performance, achieving 44.8% AP with merely 12 training epochs and 49.1% AP with 50 training epochs on COCO val 2017 with ResNet-50. Codes are available at https://github.com/ZhangGongjie/SAM-DETR .  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
0920-5691  
eISSN:  
1573-1405  
入藏号:  
WOS:001168379700001  
Indexed Date:  
2024-03-08  
Institution Address:  
Nanyang Technol Univ, Singapore, Singapore  
Institution Address:  
Carnegie Mellon Univ, Pittsburgh, PA USA  
Institution Address:  
Mohamed Bin Zayed Univ Artificial Intelligence, Abu Dhabi, U Arab Emirates  
Institution:  
Institution:  
Institution:

### 125. Lightweight Detection Network for Arbitrary-Oriented Vehicles in UAV

摘要: Recent advances in unmanned aerial vehicles (UAVs) have increased altitude capability in road-traffic monitoring. However, state-of-the-art vehicle detection methods still lack accurate abilities and lightweight structures in the UAV platform due to the background uncertainties, scales, densities, shapes, and directions of objects resulting from the UAV imagery's shooting angle. We propose a lightweight solution to detect arbitrary-oriented vehicles under uncertain backgrounds, varied resolutions, and illumination conditions. We first present a cross-stage partial bottleneck transformer (CSP BoT) module to exploit the global spatial relationship captured by multi-head self-attention, validating its implication in recessive dependencies. We then propose an angle classification prediction branch in the YOLO head network to detect arbitrarily oriented vehicles in UAV images and employ a circular smooth label (CSL) to reduce the classification loss. We further improve the multi-scale feature maps by combining the prediction head network with the adaptive spatial feature fusion block (ASFF-Head), which adapts the spatial variation of prediction uncertainties. Our method features a compact, lightweight design that automatically recognizes key geometric factors in the UAV images. It demonstrates superior performance under environmental changes while it is also easy to train and highly generalizable. This remarkable learning ability makes the proposed method applicable to geometric structure and uncertainty estimates. Extensive experiments on the UAV vehicle dataset UAV-ROD and remote sensing dataset UACS-AOD demonstrate the superiority and cost-effectiveness of the proposed method, making it practical for urban traffic and public security.  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
20  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
2  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
21  
eISSN:  
2504-446X  
入藏号:  
WOS:000801482200001  
Indexed Date:  
2022-06-06  
Institution Address:  
Chongqing Univ Posts & Telecommun, Coll Comp Sci & Technol, Chongqing 400065, Peoples R China  
Institution:

### 126. Rolling bearing fault diagnosis method using time-frequency information

摘要: Advances in deep learning methods have demonstrated remarkable development in diagnosing faults of rotating machinery. The currently popular deep neural networks suffer from design flaws in their network structure, leading to issues of long-term dependencies in fault diagnosis models built upon conventional deep neural networks. Consequently, such models exhibit insufficient global perceptual capabilities towards fault features. Furthermore, how accurately pre-trained models can diagnose faults is hugely impacted by changes in bearings' working conditions. To tackle the aforementioned issues, this study puts forth a multi-scale TransFusion (MSTF) model for diagnosing faults in rolling bearings under multiple operating conditions. Firstly, a time-frequency symmetric dot pattern transformation technique is designed to transform the original vibration signals into two-dimensional representations. This method can effectively highlight the distinctions between different fault types. Secondly, a multi-scale feature fusion module is established, which fully extracts low-level features from the time-frequency signals and reduces the complexity of the subsequent attention calculations. Meanwhile, relying on the advantages of the Transformer model in capturing global dependencies, the long-range periodic fault information is deeply mined. Finally, multi-head and multi-layer attention are visualized to enhance the interpretability of the model. After analyzing two case studies with both public and experimental datasets, the examination demonstrated that the developed model outperformed other state-of-the-art models. The diagnostic model developed in this study exhibits the ability to accurately diagnose bearing faults across multiple operating conditions while maintaining high robustness to signals contaminated with noise.  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
16  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
16  
ISSN:  
0950-7051  
eISSN:  
1872-7409  
入藏号:  
WOS:001156312500001  
Indexed Date:  
2024-02-10  
Institution Address:  
North China Elect Power Univ, Sch Energy Power & Mech Engn, Beijing 102206, Peoples R China  
Institution Address:  
Chinese Acad Sci, Inst Engn Thermophys, Beijing 100190, Peoples R China  
Institution Address:  
Liverpool John Moores Univ, Liverpool Logist Offshore & Marine LOOM Res Inst, Liverpool L3 3AF, England  
Institution Address:  
Liverpool John Moores Univ, Mech Engn & Mat Res Ctr MEMARC, Liverpool L3 3AF, England  
Institution Address:  
Shanghai Jiao Tong Univ, State Key Lab Mech Syst & Vibrat, Shanghai 200240, Peoples R China  
Institution Address:  
Guangdong Ocean Univ, Naval Architecture & Shipping Coll, Zhanjiang 524088, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 127. Automatic Extraction of Damaged Houses by Earthquake Based on Improved

摘要: Efficiently and automatically acquiring information on earthquake damage through remote sensing has posed great challenges because the classical methods of detecting houses damaged by destructive earthquakes are often both time consuming and low in accuracy. A series of deep-learning-based techniques have been developed and recent studies have demonstrated their high intelligence for automatic target extraction for natural and remote sensing images. For the detection of small artificial targets, current studies show that You Only Look Once (YOLO) has a good performance in aerial and Unmanned Aerial Vehicle (UAV) images. However, less work has been conducted on the extraction of damaged houses. In this study, we propose a YOLOv5s-ViT-BiFPN-based neural network for the detection of rural houses. Specifically, to enhance the feature information of damaged houses from the global information of the feature map, we introduce the Vision Transformer into the feature extraction network. Furthermore, regarding the scale differences for damaged houses in UAV images due to the changes in flying height, we apply the Bi-Directional Feature Pyramid Network (BiFPN) for multi-scale feature fusion to aggregate features with different resolutions and test the model. We took the 2021 Yangbi earthquake with a surface wave magnitude (Ms) of 6.4 in Yunan, China, as an example; the results show that the proposed model presents a better performance, with the average precision (AP) being increased by 9.31% and 1.23% compared to YOLOv3 and YOLOv5s, respectively, and a detection speed of 80 FPS, which is 2.96 times faster than YOLOv3. In addition, the transferability test for five other areas showed that the average accuracy was 91.23% and the total processing time was 4 min, while 100 min were needed for professional visual interpreters. The experimental results demonstrate that the YOLOv5s-ViT-BiFPN model can automatically detect damaged rural houses due to destructive earthquakes in UAV images with a good performance in terms of accuracy and timeliness, as well as being robust and transferable.  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
37  
在 BIOSIS Citation Index 中的被引频次:  
2  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
37  
eISSN:  
2072-4292  
入藏号:  
WOS:000747830800001  
Indexed Date:  
2022-02-06  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Beijing 100101, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Beijing 100049, Peoples R China  
Institution:  
Institution:

### 128. Saliency-Guided No-Reference Omnidirectional Image Quality Assessment

摘要: Due to the widespread application of the virtual reality (VR) technique, omnidirectional image (OI) has attracted remarkable attention both from academia and industry. In contrast to a natural 2-D image, an OI contains 360(omicron) x180(omicron) panoramic content, which presents great challenges for no-reference quality assessment. In this article, we propose a saliency-guided no-reference OI quality assessment (OIQA) method based on scene content understanding. Inspired by the fact that humans use hierarchical representations to grade images, we extract multiscale features from each projected viewport. Then, we integrate the texture removal and background detection techniques to obtain the corresponding saliency map of each viewport, which is subsequently utilized to guide the multiscale feature fusion from the low-level feature to the high-level one. Furthermore, motivated by the human way of understanding content, we leverage a self-attention-based Transformer to build nonlocal mutual dependencies to perceive the variations of distortion and scene in each viewport. Moreover, we also propose a content perception hypernetwork to adaptively return weights and biases for quality regressor, which is conducive to understanding the scene content and learning the perception rule for the quality assessment procedure. Comprehensive experiments validate that the proposed method can achieve competitive performances on two available databases. The code is publicly available at https://github.com/ldyorchid/SCP-OIQA.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:001346787500006  
Indexed Date:  
2024-11-24  
Institution Address:  
Anqing Normal Univ, Sch Comp & Informat, Anqing 246000, Peoples R China  
Institution Address:  
Ningbo Univ, Fac Informat Sci & Engn, Ningbo 315211, Peoples R China  
Institution Address:  
Shandong Univ Sci & Technol, Coll Elect Engn & Automat, Qingdao 266590, Peoples R China  
Institution Address:  
Anhui Normal Univ, Anhui Prov Key Lab Network & Informat Secur, Wuhu 241002, Anhui, Peoples R China  
Institution Address:  
Hangzhou Dianzi Univ, Sch Automat, Hangzhou 310061, Peoples R China  
Institution Address:  
Shanghai Univ, Sch Commun & Informat Engn, Shanghai 200444, Peoples R China  
Institution Address:  
Nanjing Univ, Sch Intelligence Sci & Technol, Nanjing 210023, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:

### 129. GFSegNet: A multi-scale segmentation model for mining area ground

摘要: Precise identification of ground fissures is of paramount importance for the safety and environmental management of coal mining areas. However, the surface environment in coal mining regions is complex, and, to date, the efficiency of artificial fissure detection has been relatively low. Therefore, we have proposed a ground fissure automatic identification model based on an encoder-decoder architecture known as the Ground Fissure Segmentation Network (GFSegNet). The encoder adopts a deep-shallow decoupled mode. The shallow network achieves spatial and spectral domain interaction by introducing adaptive Fourier convolution. The deep network adopts a hierarchical Transformer with an efficient self-attention mechanism for global modeling of fine-grained semantics. The decoder is designed as a multi-scale feature fusion structure embedded in pyramid pooling modules, aiming to efficiently utilize multi-scale ground fissure information. To advance the application of deep learning in ground fissure identification, we created a coal mining area ground fissure segmentation dataset from drone imagery, known as the mine ground fissure unmanned aerial vehicle dataset (MGF-UAV). On this dataset, the overall performance of GFSegNet surpasses the current leading segmentation models, and its reliability and generalization capabilities are further validated on additional datasets (Crack500, DeepCrack, CrackForest and ISPRS-Postdam). This research has brought expansion and innovation to the field of automatic ground fissure recognition in coal mining areas, offering new perspectives and methodologies for the application of deep learning techniques in this domain.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
6  
被引频次合计:  
6  
ISSN:  
1569-8432  
eISSN:  
1872-826X  
入藏号:  
WOS:001216401100001  
Indexed Date:  
2024-05-14  
Institution Address:  
China Univ Min & Technol Beijing, Coll Geosci & Surveying Engn, Beijing 100083, Peoples R China  
Institution Address:  
State Key Lab Coal Min & Clean Utilizat, Beijing 100013, Peoples R China  
Institution Address:  
Minist Nat Resources, Key Lab Mine Ecol Effects & Systemat Restorat, Beijing 100081, Peoples R China  
Institution:  
Institution:  
State Key Lab Coal Min & Clean Utilizat  
Institution:

### 130. A lightweight spatially-aware classification model for breast cancer

摘要: Breast cancer is a prevalent malignant tumour with high global incidence. Its diagnosis relies primarily on the analysis of pathological breast images. Owing to the complex organisation of the tumour microenvironment, neural network models are essential as efficient classification tools in the field of pathological image analysis. This study introduced spatially-aware attention swift parallel convolution network (SPA-SPCNet), a lightweight and low-latency model for classifying breast pathologies. A novel module for multi-scale feature extraction was constructed using a depthwise separable convolution method. It focuses on the multi-scale features of pathological images to alleviate recognition problems caused by similar local features in breast cancer tissues. The module concatenates the convolutions of different kernels from three branches. Second, a lightweight dynamic spatially-aware attention module was introduced to integrate the visual graph convolutional architecture in a branch. This allowed the model to capture the spatial structure and relationships in image, enabling better handling of the unique spatial distribution relationship between breast cancer tissue structures. The other branch utilises a self-attention mechanism in the transformer. The module can dynamically adjust the attention of the model to different regions in the image, allowing it to focus on the key features of the complex spatial distribution of breast cancer tissue. This feature fusion method enabled the model to capture both global semantics and local details. Compared with existing lightweight models, the proposed model has advantages in terms of tissue structure classification accuracy, parameter quantity, floating-point operations, and real-time inference speed, providing a powerful tool for computer-aided breast pathological image classification.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0208-5216  
入藏号:  
WOS:001303162900001  
Indexed Date:  
2024-09-07  
Institution Address:  
Shandong Univ Tradit Chinese Med, Coll Intelligence & Informat Engn, Jinan 250355, Shandong, Peoples R China  
Institution:

### 131. ADT-UNet: An Innovative Algorithm for Glioma Segmentation in MR Images

摘要: The precise delineation of glioma tumors is of paramount importance for surgical and radiotherapy planning. Presently, the primary drawbacks associated with the manual segmentation approach are its laboriousness and inefficiency. In order to tackle these challenges, a deep learning-based automatic segmentation technique was introduced to enhance the efficiency of the segmentation process. We proposed ADT-UNet, an innovative algorithm for segmenting glioma tumors in MR images. ADT-UNet leveraged attention-dense blocks and Transformer as its foundational elements. It extended the U-Net framework by incorporating the dense connection structure and attention mechanism. Additionally, a Transformer structure was introduced at the end of the encoder. Furthermore, a novel attention-guided multi-scale feature fusion module was integrated into the decoder. To enhance network stability during training, a loss function was devised that combines Dice loss and binary cross-entropy loss, effectively guiding the network optimization process. On the test set, the DSC was 0.933, the IOU was 0.878, the PPV was 0.942, and the Sen was 0.938. Ablation experiments conclusively demonstrated that the inclusion of all the three proposed modules led to enhanced segmentation accuracy within the model. The most favorable outcomes were observed when all the three modules were employed simultaneously. The proposed methodology exhibited substantial competitiveness across various evaluation indices, with the three additional modules synergistically complementing each other to collectively enhance the segmentation accuracy of the model. Consequently, it is anticipated that this method will serve as a robust tool for assisting clinicians in auxiliary diagnosis and contribute to the advancement of medical intelligence technology.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0899-9457  
eISSN:  
1098-1098  
入藏号:  
WOS:001282286500001  
Indexed Date:  
2024-08-09  
Institution Address:  
Nanjing Univ Chinese Med, Taizhou Affiliated Hosp, Taizhou, Jiangsu, Peoples R China  
Institution Address:  
Xuzhou Med Univ, Sch Med Informat & Engn, Xuzhou, Jiangsu, Peoples R China  
Institution Address:  
Nanjing Univ Chinese Med, Nanjing, Jiangsu, Peoples R China  
Institution Address:  
Nanjing Med Univ, Affiliated Taizhou Peoples Hosp, Dept Radiol, Taizhou, Jiangsu, Peoples R China  
Institution Address:  
Haian Hosp Tradit Chinese Med, Dept Gen Surg, Nantong, Jiangsu, Peoples R China  
Institution Address:  
Nanjing Med Univ, Affiliated Taizhou Peoples Hosp, Taizhou Sch Clin Med, Dept Endocrinol, Taizhou, Jiangsu, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Haian Hosp Tradit Chinese Med  
Institution:

### 132. U-Shape Transformer for Underwater Image Enhancement

摘要: The light absorption and scattering of underwater impurities lead to poor underwater imaging quality. The existing data-driven based underwater image enhancement (UIE) techniques suffer from the lack of a large-scale dataset containing various underwater scenes and high-fidelity reference images. Besides, the inconsistent attenuation in different color channels and space areas is not fully considered for boosted enhancement. In this work, we built a large scale underwater image (LSUI) dataset, which covers more abundant underwater scenes and better visual quality reference images than existing underwater datasets. The dataset contains 4279 real-world underwater image groups, in which each raw image's clear reference images, semantic segmentation map and medium transmission map are paired correspondingly. We also reported an U-shape Transformer network where the transformer model is for the first time introduced to the UIE task. The U-shape Transformer is integrated with a channel-wise multi-scale feature fusion transformer (CMSFFT) module and a spatial-wise global feature modeling transformer (SGFMT) module specially designed for UIE task, which reinforce the network's attention to the color channels and space areas with more serious attenuation. Meanwhile, in order to further improve the contrast and saturation, a novel loss function combining RGB, LAB and LCH color spaces is designed following the human vision principle. The extensive experiments on available datasets validate the state-of-the-art performance of the reported technique with more than 2dB superiority. The dataset and demo code are available at https://bianlab.github.io/.  
在 BIOSIS Citation Index 中的被引频次:  
10  
在中国科学引文数据库中的被引频次:  
10  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
228  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
238  
ISSN:  
1057-7149  
eISSN:  
1941-0042  
入藏号:  
WOS:001001353000005  
PubMed ID:  
37200123  
Indexed Date:  
2023-06-27  
Institution Address:  
Beijing Inst Technol, MIIT Key Lab Complex Field Intelligent Sensing, Beijing 100081, Peoples R China  
Institution Address:  
Beijing Inst Technol Jiaxing, Yangtze Delta Reg Acad, Jiaxing 314019, Peoples R China  
Institution:  
Institution:  
Beijing Inst Technol Jiaxing

### 133. Research on Improved MobileViT Image Tamper Localization Model

摘要: As image manipulation technology advances rapidly, the malicious use of image tampering has alarmingly escalated, posing a significant threat to social stability.In the realm of image tampering localization, accurately localizing limited samples, multiple types, and various sizes of regions remains a multitude of challenges. These issues impede the model's universality and generalization capability and detrimentally affect its performance. To tackle these issues, we propose FL-MobileViT-an improved MobileViT model devised for image tampering localization. Our proposed model utilizes a dual-stream architecture that independently processes the RGB and noise domain, and captures richer traces of tampering through dual-stream integration. Meanwhile, the model incorporating the Focused Linear Attention mechanism within the lightweight network (MobileViT). This substitution significantly diminishes computational complexity and resolves homogeneity problems associated with traditional Transformer attention mechanisms, enhancing feature extraction diversity and improving the model's localization performance. To comprehensively fuse the generated results from both feature extractors, we introduce the ASPP architecture for multi-scale feature fusion. This facilitates a more precise localization of tampered regions of various sizes. Furthermore, to bolster the model's generalization ability, we adopt a contrastive learning method and devise a joint optimization training strategy that leverages fused features and captures the disparities in feature distribution in tampered images. This strategy enables the learning of contrastive loss at various stages of the feature extractor and employs it as an additional constraint condition in conjunction with cross-entropy loss. As a result, overfitting issues are effectively alleviated, and the differentiation between tampered and untampered regions is enhanced. Experimental evaluations on five benchmark datasets (IMD-20, CASIA, NIST-16, Columbia and Coverage) validate the effectiveness of our proposed model. The meticulously calibrated FL-MobileViT model consistently outperforms numerous existing general models regarding localization accuracy across diverse datasets, demonstrating superior adaptability.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1546-2218  
eISSN:  
1546-2226  
入藏号:  
WOS:001315133800011  
Indexed Date:  
2024-09-23  
Institution Address:  
Xian Univ Posts & Telecommun, Sch Comp Sci & Technol, Xian 710121, Peoples R China  
Institution Address:  
Xian Univ Posts & Telecommun, Shaanxi Key Lab Network Data Anal & Intelligent Pr, Xian 710121, Peoples R China  
Institution:  
Institution:

### 134. CFNet: Cross-scale fusion network for medical image segmentation

摘要: Learning multi-scale feature representations is essential for medical image segmentation. Most existing frameworks are based on U-shape architecture in which the high-resolution representation is recovered progressively by connecting different levels of the decoder with the low-resolution representation from the encoder. However, intrinsic defects in complementary feature fusion inhibit the U-shape from aggregating efficient global and discriminative features along object boundaries. While Transformer can help model the global features, their computation complexity limits the application in real-time medical scenarios. To address these issues, we propose a Cross-scale Fusion Network (CFNet), combining a cross-scale attention module and pyramidal module to fuse multi-stage/global context information. Specifically, we first utilize large kernel convolution to design the basic building block capable of extracting global and local information. Then, we propose a Bidirectional Atrous Spatial Pyramid Pooling (BiASPP), which employs atrous convolution in the bidirectional paths to capture various shapes and sizes of brain tumors. Furthermore, we introduce a cross- stage attention mechanism to reduce redundant information when merging features from two stages with different semantics. Extensive evaluation was performed on five medical image segmentation datasets: a 3D volumetric dataset, namely Brats benchmarks. CFNet-L achieves 85.74% and 90.98% dice score for Enhanced Tumor and Whole Tumor on Brats2018, respectively. Furthermore, our largest model CFNet-L outperformed other methods on 2D medical image. It achieved 71.95%, 82.79%, and 80.79% SE for STARE, DRIVE, and CHASEDB1, respectively. The code will be available at https://github.com/aminabenabid/CFNet  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1319-1578  
eISSN:  
2213-1248  
入藏号:  
WOS:001278512800001  
Indexed Date:  
2024-08-04  
Institution Address:  
Zhejiang Normal Univ, Coll Math Med, 688 Yingbin Rd, Jinhua 321004, Zhejiang, Peoples R China  
Institution Address:  
Zhejiang Normal Univ, Coll Comp Sci & Technol, 688 Yingbin Rd, Jinhua 321004, Zhejiang, Peoples R China  
Institution Address:  
Zhejiang Normal Univ, Zhejiang Inst Photoelect, 688 Yingbin Rd, Jinhua 321004, Zhejiang, Peoples R China  
Institution Address:  
Zhejiang Normal Univ, Zhejiang Inst Adv Light Source, 688 Yingbin Rd, Jinhua 321004, Zhejiang, Peoples R China  
Institution Address:  
Univ Constantine 2 Abdelhamid Mehri, Fac New Technol Informat & Commun, Constantine 25000, Algeria  
Institution:  
Institution:  
Institution:  
Institution:  
Institution:  
Univ Constantine 2 Abdelhamid Mehri

### 135. Tropical Rice Mapping Using Time-Series SAR Images and ESF-Seg Model in

摘要: Tropical and subtropical Asia is the major rice-producing region in the world, but the complexity of the cropping system and the diversity of the topography bring challenges to the accurate monitoring of rice cultivation. To address this difficulty, a new deep learning model, ESF-Seg, is proposed in this study to extract the annual tropical rice distribution using monthly averaged time-series Sentinel-1 VH data. The ESF-Seg adopts the Efficient Adaptive Sparse Transformer (EAT) to remove redundant information from input features. The Channel Attention Bridge Block (CAB) and Spatial Attention Bridge Block (SAB) modules are introduced to refine the information. Meanwhile, with the FreqFusion-KAN (FreqK) module, the loss of information can be reduced through the multi-scale feature fusion strategy. The proposed method is evaluated in the Hainan Province of China, an important tropical arable zone with diverse crop resources and complicated croplands. First, ablation experiments are conducted. Compared to the classical SegFormer model, the ESF-Seg model improves on the mIOU by 4.99% and on the mPA by 2.65%. Subsequently, compared to the RF, U-Net, and the original SegFormer model, the overall accuracy (OA) of the ESF-Seg model on the validation samples increased by 11.02%, 2.01%, and 1.33%, and the F1 score improved by 0.0756, 0.0624, and 0.0490, reaching 98.31% and 0.9506, respectively. Furthermore, products showing the annual rice distribution from 2019 to 2023 in Hainan are generated, which exhibit good alignments with the statistical data, surpassing other existing products with an RMSE of 5.4004 Kha. As indicated by the rice mapping products, the proposed method preserves the integrity of the rice parcels in the fragmented croplands, thus providing a new opportunity for the continuous monitoring of tropical rice distribution with high accuracy.  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
eISSN:  
2072-4292  
入藏号:  
WOS:001404701800001  
Indexed Date:  
2025-02-01  
Institution Address:  
Hainan Aerosp Informat Res Inst, Key Lab Earth Observat Hainan Prov, Sanya 572029, Peoples R China  
Institution Address:  
Chinese Acad Sci, Aerosp Informat Res Inst, Key Lab Digital Earth Sci, Beijing 100094, Peoples R China  
Institution Address:  
Int Res Ctr Big Data Sustainable Dev Goals, Beijing 100094, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Coll Resources & Environm, Beijing 100049, Peoples R China  
Institution:  
Hainan Aerosp Informat Res Inst  
Institution:  
Institution:  
Institution:

### 136. IAE-SDNet: An End-to-End Image Adaptive Enhancement and Wheat Scab

摘要: Wheat scab is a severe fungal disease caused by Fusarium. The precise detection of scab is crucial for improving the wheat yield. The utilization of RGB images captured by unmanned aerial vehicle (UAV) can effectively enhance the wheat scab detection efficiency. However, UAV images are influenced by environmental conditions. In addition, the scab in the image is extremely small and the background is complex. Therefore, to boost the wheat scab detection accuracy, this study proposes an end-to-end image adaptive enhancement and wheat scab detection network (IAE-SDNet). In the head of the network, an image adaptive enhancement (IAE) module aims to enhance the quality of UAV images and performs end-to-end learning with the transformer-based object detection module SDNet. A cascading inverse residual module (CIRMB) is designed to boost the feature extraction capacity for the disease area. A deformable multihead attention encoder (DMHA-Encoder) is deployed to augment the semantic information of the advanced features of wheat while maintaining the representation ability of global attention. A multiscale feature fusion block (MFFB) is designed to effectively fuse wheat disease features of different scales, reducing the impact caused by the inconsistency in size between the complex field background and the disease area. Finally, to optimize the detection of small lesion areas, an MNIoU bounding-box regression loss function is proposed. The experiment indicates that this method has improved average precision (AP) by 6% and the Recall by 5% compared with the baseline network. This study provides a feasible approach for the precise detection of wheat scab, with high application prospects.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0196-2892  
eISSN:  
1558-0644  
入藏号:  
WOS:001358693100024  
Indexed Date:  
2024-12-05  
Institution Address:  
Anhui Univ, Natl Engn Res Ctr Agroecol Big Data Anal & Applica, Hefei 230601, Peoples R China  
Institution Address:  
CNTC, Zhengzhou Tobacco Res Inst, Zhengzhou 450001, Peoples R China  
Institution Address:  
Chinese Acad Sci, Hefei Inst Phys Sci, Hefei 230031, Peoples R China  
Institution:  
Institution:  
Institution:

### 137. Multi-scale feature map fusion encoding for underwater object

摘要: Underwater object segmentation presents significant challenges due to the degradation of image quality and the complexity of underwater environments. In recent years, deep learning has provided an effective approach for object segmentation. However, DeepLabV3+, as a classical model for general scenes, shows limitations in achieving accurate and real-time segmentation in complex underwater conditions. To address this issue, we propose a DeepLab-FusionNet, an extended version of DeepLabV3+, specifically designed for underwater object segmentation. The model utilizes a multi-resolution parallel branch structure to extract multi-scale information and employs an improved inverted residual structure as the basic feature extraction module in the encoding network. Structural reparameterization technique is introduced to optimize inference speed and memory access costs during the inference stage. Additionally, a module for linking deep and shallow level information is constructed to reduce the loss of detail and spatial information during downsampling and convolution. Evaluation on the SUIM dataset shows a 3.3% increase in mean Intersection over Union (mIoU) and a speed improvement of 34 frames per second (FPS) compared to the baseline model DeepLabV3+. Further comparisons with other classic lightweight models and Transformer-based models on the UIIS and TrashCan datasets demonstrate that our model achieves good accuracy and balanced computational efficiency in challenging underwater environments. Although there is room for improvement due to overfitting and fixed convolution kernel limitations, future integration with Transformer methods is planned. Our model offers an effective solution for real-time target segmentation for underwater robots, with broad applications in human exploration and development of marine resources. Our codes are available at: https://github.com/sunmer1rain/deeplabv\_fusionnet  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:001378128600007  
Indexed Date:  
2024-12-25  
Institution Address:  
Shenzhen Univ, Coll Mechatron & Control Engn, Shenzhen 518060, Peoples R China  
Institution Address:  
Shenzhen Univ, Coll Phys & Optoelect Engn, Shenzhen 518060, Peoples R China  
Institution:  
Institution:

### 138. Dual-Branch Multiscale Optimization Network for Enhancing Low-Light

摘要: The existing deep learning-based low-light image enhancement models have demonstrated their validity on many benchmark datasets. However, it is not easy to enhance the brightness, color, contrast, and other information of images while maintaining the quality of image details with these models. To enhance the visual perception of images under low-light conditions at night, improve driving visibility, and increase the accuracy of obstacle detection and segmentation, a dual-branch multiscale optimization (DBMO) network is proposed in this article. The global branch employs a multiscale network based on the transformer, which preserves high-resolution input images while effectively integrating multiscale features to capture global image information. Meanwhile, the detailed branch introduces an adjustment denoising network based on wavelet transform. It performs noise suppression and detail enhancement on the high- and low-frequency information obtained from wavelet decomposition, thereby enhancing image details while balancing brightness and contrast. Finally, the feature information extracted from both branches is adaptively weighted and fused to produce the final enhanced image. The experimental results on a real rail transit obstacle dataset demonstrate that DBMO significantly improves images' overall brightness and color balance and achieves the highest accuracy improvement in rail transit obstacle detection and segmentation. Compared to baseline models YOLO-v8 and Deeplab-v3+, obstacle target detection and semantic segmentation accuracy improved by 3.6% and 7.65%, respectively. This model can be applied to nighttime train assistance systems and shows promising potential in obstacle detection and segmentation tasks.  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1530-437X  
eISSN:  
1558-1748  
入藏号:  
WOS:001418812500012  
Indexed Date:  
2025-02-20  
Institution Address:  
Guangxi Univ, Sch Mech Engn, Guangxi Mfg Syst & Adv Mfg Technol Key Lab, Nanning 530004, Peoples R China  
Institution:

### 139. MG-Net: A fetal brain tissue segmentation method based on multiscale

摘要: Background and Objective: Fetal brain tissue segmentation provides foundational support for comprehensively understanding the neurodevelopment of normal and congenital disease-affected fetuses. Manual labeling is very time-consuming, and automated segmentation methods can greatly improve the efficiency of doctors. At the same time, fetal brain tissue undergoes various changes throughout the pregnancy, leading to a continuous change in tissue contrast, which greatly increases the difficulty of training segmentation methods. This study aims to develop an automated segmentation model that can efficiently and accurately segment fetal brain tissue, improving the workflow for medical professionals. Methods: We propose a novel deep learning-based segmentation model that incorporates three innovative components: Firstly, a new Dual Dilated Attention Block (DDAB) is proposed in the encoder part to enhance the feature extraction of local spatial and structural contextual information. Secondly, a Multi-scale Deformable Transformer (MSDT) is integrated into the bottleneck to improve the feature extraction of global information on local spatial and structural contextual information. Thirdly, we use a novel block based on Graph Convolution Attention (GCAB) in the decoder, which effectively enhances the features at the decoder.The code is available at https://github.com/unicoco7/MG-Net/. Results: We trained and tested on the FeTA 2021 and FeTA 2022 datasets, and evaluated using seven popular metrics, including Dice, IoU, MAE, BoundaryF, PRE, SEN, and SPE. Compared to the current state-of-the-art 3D segmentation models such as nnFormer, SwinUNETR, and 3DUX-net, our proposed method has surpassed all of them in metrics like Dice, IoU, and MAE. Specifically, on the FeTA 2021 dataset, our model achieved a Dice of 0.8666, an IoU of 0.7646, and an MAE of 0.0027; on the FeTA 2022 dataset, it achieved a Dice of 0.8552, an IoU of 0.7470, and an MAE of 0.0005. Conclusion: In this paper, we propose a model for three-dimensional fetal brain tissue segmentation based on multi-scale feature fusion and graph convolution attention mechanism, and conduct experimental evaluation on the FeTA 2021 and FeTA 2022 datasets. Understanding the boundaries of fetal brain tissue is crucial for doctors' diagnosis, so the proposed model is expected to improve the speed and accuracy of doctors' diagnoses.  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
0169-2607  
eISSN:  
1872-7565  
入藏号:  
WOS:001335112700001  
PubMed ID:  
39395303  
Indexed Date:  
2024-10-25  
Institution Address:  
Chinese Acad Sci, Res Ctr Med AI, Shenzhen Inst Adv Technol, Shenzhen 518055, Peoples R China  
Institution Address:  
Univ Chinese Acad Sci, Dept Artificial Intelligence, Beijing 100049, Peoples R China  
Institution Address:  
Nanjing Univ, Nanjing Drum Tower Hosp, Affiliated Hosp, Dept Radiol,Med Sch, Nanjing, Peoples R China  
Institution:  
Institution:  
Institution:

### 140. Regional perception and multi-scale feature fusion network for cardiac

摘要: Objective. Cardiovascular disease (CVD) is a group of diseases affecting cardiac and blood vessels, and short-axis cardiac magnetic resonance (CMR) images are considered the gold standard for the diagnosis and assessment of CVD. In CMR images, accurate segmentation of cardiac structures (e.g. left ventricle) assists in the parametric quantification of cardiac function. However, the dynamic beating of the heart renders the location of the heart with respect to other tissues difficult to resolve, and the myocardium and its surrounding tissues are similar in grayscale. This makes it challenging to accurately segment the cardiac images. Our goal is to develop a more accurate CMR image segmentation approach. Approach. In this study, we propose a regional perception and multi-scale feature fusion network (RMFNet) for CMR image segmentation. We design two regional perception modules, a window selection transformer (WST) module and a grid extraction transformer (GET) module. The WST module introduces a window selection block to adaptively select the window of interest to perceive information, and a windowed transformer block to enhance global information extraction within each feature window. The WST module enhances the network performance by improving the window of interest. The GET module grids the feature maps to decrease the redundant information in the feature maps and enhances the extraction of latent feature information of the network. The RMFNet further introduces a novel multi-scale feature extraction module to improve the ability to retain detailed information. Main results. The RMFNet is validated with experiments on three cardiac data sets. The results show that the RMFNet outperforms other advanced methods in overall performance. The RMFNet is further validated for generalizability on a multi-organ data set. The results also show that the RMFNet surpasses other comparison methods. Significance. Accurate medical image segmentation can reduce the stress of radiologists and play an important role in image-guided clinical procedures.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
6  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
6  
ISSN:  
0031-9155  
eISSN:  
1361-6560  
入藏号:  
WOS:000980289300001  
PubMed ID:  
36958057  
Indexed Date:  
2023-05-14  
Institution Address:  
Hebei Univ Technol, Sch Artificial Intelligence, Tianjin 300401, Peoples R China  
Institution Address:  
Hebei Univ Technol, Sch Elect & Informat Engn, Tianjin 300401, Peoples R China  
Institution Address:  
Univ Massachusetts Lowell, Dept Elect & Comp Engn, Lowell, MA 01854 USA  
Institution:  
Institution:  
Institution:

### 141. Region segmentation and localization method of water plants in crab pond

摘要: At present, the cleaning of water plants in crab pond is primarily relied on manual operations, which is characterized by high labor intensity and low efficiency. The automatic water plant cleaning boat can greatly reduce the labor intensity and improve the operation efficiency, and determining the water plant distribution in crab pond is the key basis for planning the efficient operation path of the cleaning boat. Addressing the challenges posed by the high similarity between water plants and shoreline vegetation, as observed in drone-captured images of the crab pond, which complicates the accurate distinction between them, a region segmentation and localization method for water plants in crab pond based on improved YOLOv8n-seg was proposed in this paper. The focus of the improved model lies in the reduction of the model size and the enhancement of the recall rate. Firstly, in order to diminish the model size, the lightweight HGNetv2 (hierarchical graph network) based on RTDETR (real-time detection transformer) was used as the backbone feature extraction network. Secondly, the neck network was reconstructed based on the lightweight structure of Efficient Rep to reduce the number of parameters and enhance the multiscale feature fusion ability of the mode. Finally, the SegNext attention mechanism was introduced in the feature extraction layer to enhance the sensitivity of the model to the water plant area. In order to eliminate the redundant regions generated by the model in the recognition process and further improve the segmentation accuracy, the binarization processing was used to optimize the segmentation results, and the image processing algorithm was combined to screen the area of the water plant area. After coordinate conversion, the precise longitude and latitude coordinates of water plant contour were obtained. The experimental results showed that the improved model exhibited a strong discriminatory and segmentation effect on water plants in crab ponds. The parameter number, calculation quantity and model size of the improved model were only 1.49 M, 8.4 GFLOPs, and 3.27 MB, respectively. Compared with the original YOLOv8n-seg model, the parameter number, calculation quantity and model size of the improved model were reduced by 54.3%, 30.6% and 49.7%, respectively. The model before and after improvement was deployed to the Jetson Orin Nano embedded AI computer for testing. It was found that the preprocessing speed of the improved model was almost the same as that of the YOLOv8n-seg model. The inference speed, post processing speed and segment speed were increased by 30.6%, 46.4% and 32.3%, respectively. In addition, the recall, precision, and mean average precision of the improved model achieved values of 91.5%, 89.3% and 95.6%, respectively. The improved model achieved the best balance of calculation, parameters, recall and detection accuracy comparing with YOLOv5sseg, YOLOv8s-SwinTransformer, YOLOv8s-seg and YOLOv8n-GoldYOLO models. The coordinate conversion test showed that the minimum distance error of water plant location accuracy was 0.13m, the maximum distance error was 0.33 m, and the average distance error was 0.22 m, which verified that the improved model can meet the requirements of water plant region segmentation and localization in crab pond. The finding can provide an important reference for the automatic operation path planning of water plant cleaning boat.  
摘要:  
目前蟹塘内水草清理以人工作业为主,劳动强度大,作业效率低,自动水草清理船能大幅降低劳动强度,提高作业效率,而确定蟹塘水草分布是规划水草清理船高效作业路径的关键基础。针对无人机采集的蟹塘水草图像中水草与岸边植被相似性高、难以准确区分等问题,该研究提出一种基于改进YOLOv8n-seg的蟹塘水草区域分割与定位方法。改进模型从降低模型大小和提高召回率的角度出发,首先基于RT-DETR(real-time detection transformer)的HGNetv2(hierarchical graph network)轻量化网络结构重设主干特征提取网络,缩减模型体积;其次以Efficient Rep轻量化结构为基准重构颈部网络,在降低参数量的同时增强模型的多尺度特征融合能力;接着在特征提取层引入SegNext注意力机制,加强模型对蟹塘水草区域的敏感度。为了消除模型在识别过程中产生的冗余区域,进一步提高分割精度,采用二值化处理对分割结果进行优化,并结合图像处理算法对水草区域进行面积筛选;经过坐标转换后得到精确水草轮廓经纬度坐标。试验结果表明:改进模型对蟹塘水草具有良好的区分度和分割效果,其参数量和计算量分别为1.49 M和8.4 GFLOPs,召回率和平均精度均值分别为91.5%和95.6%,与原YOLOv8n-seg模型相比,模型体积减小了49.7%,分割速度提升了32.3%。在坐标转换试验中,水草定位精度平均误差为0.22 m,验证了改进模型能够满足蟹塘水草区域分割与定位要求。研究结果为后续水草清理船自动作业路径规划研究提供参考。  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1002-6819  
入藏号:  
CSCD:7828658  
Indexed Date:  
2025-01-07  
Institution Address:  
江苏大学电气信息工程学院;;常州东风农机集团有限公司, ;;, 镇江;;常州, ;; 212013;;213200  
Institution Address:  
江苏大学电气信息工程学院, 镇江, 江苏 212013, 中国  
Institution Address:  
江苏大学电气信息工程学院;;江苏大学, ;;江苏省机械工业设施农业测控技术与装备重点实验室, 镇江;;镇江, ;; 212013;;212013  
Institution Address:  
College of Electrical and Information Engineering,Jiangsu University;;Changzhou Dongfeng Agricultural Machinery Group Co.,LTD., ;;, Zhenjiang;;Changzhou, ;; 212013;;213200  
Institution Address:  
College of Electrical and Information Engineering,Jiangsu University, Zhenjiang, Jiangsu 212013, China  
Institution Address:  
College of Electrical and Information Engineering,Jiangsu University;;Jiangsu University, ;;Key Laboratory of Agricultural Measurement and Control Technology and Equipment for Machinery Industry, Zhenjiang;;Zhenjiang, ;; 212013;;212013  
Institution:  
江苏大学电气信息工程学院;;常州东风农机集团有限公司  
Institution:  
江苏大学电气信息工程学院  
Institution:  
江苏大学电气信息工程学院;;江苏大学  
Institution:  
College of Electrical and Information Engineering,Jiangsu University;;Changzhou Dongfeng Agricultural Machinery Group Co.,LTD.  
Institution:  
College of Electrical and Information Engineering,Jiangsu University  
Institution:  
College of Electrical and Information Engineering,Jiangsu University;;Jiangsu University

### 142. Rotating box multi-objective visual tracking algorithm for vibration

摘要: Visual displacement measurement methods for flexible structural bodies like large-span bridges has gained wide popularity in recent years, but practical applications still have some limitations. For instance, when acquiring images of large-span flexible bridges at a distance, the slight angular tilt of the detection target due to irregular vibrations can cause extremely serious misfit errors in the displacement curves returned by the vision measurement algorithm. To improve the reliability of vibration displacement measurement of flexible structural bodies, this paper takes the bridge subjected to external excitation in the acquired image sequence as the object of vibration displacement measurement and uses a designed high-precision displacement measurement algorithm for a single-stage rotating target tracking anchor-free box to track the vibration displacement of the target in the flexible structural body. We first extract multi-scale feature information of bridge model image sequences using the improved YOLOv5-s backbone network and combine the Transformer self-attention mechanism with PANet to perform a top-down and bottom-up bi-directional fusion of target feature maps at three different scales to achieve semantic feature fusion of shallow and deep information. Second, the improved Efficient Decoupled Head performs the detection of rotating target centroid offset and bounding box size. Finally, the detected results are passed into the multi-objective tracking algorithm ByteTrack, which strengthens the spatio-temporal correlation between frames and obtains a better-fitting vibration displacement curve. The validation and comparison of traditional visual measurement methods and deep learning measurement methods on cable-stayed bridge models, small arch bridges, and large span bridges show that the vibration displacement trajectories regressed by the algorithm in this paper have the best fit with the actual vibration displacement trajectories, which also verifies that the algorithm in this paper has good potential for engineering applications and implementation space in the field of condition monitoring of flexible structural bodies.  
Web of Science 核心合集中的 "被引频次":  
13  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
13  
ISSN:  
0888-3270  
eISSN:  
1096-1216  
入藏号:  
WOS:001050753700001  
Indexed Date:  
2023-08-29  
Institution Address:  
Kunming Univ Sci & Technol, Fac Mech & Elect Engn, Kunming 650500, Peoples R China  
Institution Address:  
Lab Adv Equipment Intelligent Mfg Technol Yunnan P, Kunming 650500, Peoples R China  
Institution:  
Institution:  
Lab Adv Equipment Intelligent Mfg Technol Yunnan P

### 143. Pulp grade monitoring using binocular image through multi-scale feature

摘要: Pulp grade is an important indicator for performance monitoring in froth flotation process. Previous studies have shown that the pulp grade can be predicted more accurately by using the stereo vision information of froth images. However, due to the low intra-class variation among froth images under the identical working conditions and the similarity in shape and texture among bubbles, it is challenging for current binocular image-based methods for grade prediction. Therefore, to accurately predict the key performance indicators in flotation process, a prediction model based on binocular image fusion is proposed in this paper. First, a calculation method of froth image saliency map is proposed, and the saliency map is used as a priori knowledge to guide the prediction model to learn the characteristics of the region of interest in the image. This measure aims to solve the problem of difficulty in grade prediction caused by low intra-class differences in froth images. Then, a multi-scale feature cross-attention fusion network is introduced, wherein the multi-scale features of the left and right views serve as attention mechanism gating signals to extract common feature from binocular image. After that, a pulp grade prediction model is developed based on Video Transformer Network. The results on actual industrial flotation datasets show that, compared with other pulp grade prediction methods, our proposed approach reduces the mean absolute error and root mean square error by 22.84% and 23.55%, respectively.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
1474-0346  
eISSN:  
1873-5320  
入藏号:  
WOS:001301246500001  
Indexed Date:  
2024-09-04  
Institution Address:  
Cent South Univ, Sch Automat, Changsha 410083, Hunan, Peoples R China  
Institution Address:  
Cent South Univ, Sch Math & Stat, Changsha 410083, Hunan, Peoples R China  
Institution Address:  
Changsha Univ, Coll Comp Sci & Engn, Changsha 410022, Hunan, Peoples R China  
Institution Address:  
Cent South Univ, Dundee Int Inst, Changsha 410083, Hunan, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:

### 144. PEDNet: A Lightweight Detection Network of Power Equipment in Infrared

摘要: As a promising and noncontact detection technique, machine vision has been widely used in fault diagnosis of substation equipment. The rapid and accurate detection of substation equipment in infrared images is one of the key steps for automatic fault diagnosis. However, the complexity of image background, the low contrast of infrared images, and the rotational targets in infrared images pose a great challenge to detection task. This study aims to improve the detection accuracy of the model while having real-time detection speed and propose a lightweight power equipment detection network (PEDNet) based on You Only Look Once (YOLOv4)-tiny. First, a novel global information aggregation module (GIAM) is constructed to guide the network to focus on the salient regions where the target equipment is located. Second, an improved spatial transformer network (ISTN) is introduced to reduce the impact of rotational targets on detection accuracy. Finally, a feature enhanced fusion network (FEFN) is designed through the use of a multiscale feature cross-fusion structure. It can fully fuse the feature information of the salient region, the rotational targets, and the strong semantic information. The experimental results show that the proposed PEDNet can reach 92.66% detection accuracy and 107.07 frames/s real-time detection speed on the testing datasets. Compared with YOLOv4-tiny, there is a small sacrifice in detection speed, but the detection accuracy is improved and significantly higher than the existing state-of-the-art (SOTA) object detection models.  
在中国科学引文数据库中的被引频次:  
2  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
32  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
34  
ISSN:  
0018-9456  
eISSN:  
1557-9662  
入藏号:  
WOS:000922147200118  
Indexed Date:  
2024-06-11  
Institution Address:  
Xiangtan Univ, Coll Automat & Elect Informat, Xiangtan 411105, Peoples R China  
Institution Address:  
Hunan Univ Arts & Sci, Coll Comp & Elect Engn, Changde 415000, Peoples R China  
Institution Address:  
Xiangtan Univ, Key Lab Hunan Prov Internet Things & Informat Sec, Xiangtan 411105, Peoples R China  
Institution Address:  
Hunan Acad Bldg Res Co Ltd, Changsha 410014, Peoples R China  
Institution Address:  
Hunan Prov Key Lab Control Technol Distributed El, Changde 415000, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Hunan Acad Bldg Res Co Ltd  
Institution:  
Hunan Prov Key Lab Control Technol Distributed El

## Support Vector Machine (SVM)

### 1. Multiscale Analysis of Alzheimer's Disease Using Feature Fusion in

摘要: Introduction: This research is focused on early detection of Alzheimer's disease (AD) using a multiscale feature fusion framework, combining biomarkers from memory, vision, and speech regions extracted from magnetic resonance imaging and positron emission tomography images.  
Methods: Using 2D gray level co-occurrence matrix (2D-GLCM) texture features, volume, standardized uptake value ratios (SUVR), and obesity from different neuroimaging modalities, the study applies various classifiers, demonstrating a feature importance analysis in each region of interest. The research employs four classifiers, namely linear support vector machine, linear discriminant analysis, logistic regression (LR), and logistic regression with stochastic gradient descent (LRSGD) classifiers, to determine feature importance, leading to subsequent validation using a probabilistic neural network classifier.  
Results: The research highlights the critical role of brain texture features, particularly in memory regions, for AD detection. Significant sex-specific differences are observed, with males showing significance in texture features in memory regions, volume in vision regions, and SUVR in speech regions, while females exhibit significance in texture features in memory and speech regions, and SUVR in vision regions. Additionally, the study analyzes how obesity affects features used in AD prediction models, clarifying its effects on speech and vision regions, particularly brain volume.  
Conclusion: The findings contribute valuable insights into the effectiveness of feature fusion, sex-specific differences, and the impact of obesity on AD-related biomarkers, paving the way for future research in early AD detection strategies and cognitive impairment classification.

### 2. Recognition of honeycomb lung in CT images based on improved MobileNet

摘要: Purpose The research is to improve the efficiency and accuracy of recognition of honeycomb lung in CT images. Methods Deep learning methods are used to achieve automatic recognition of honeycomb lung in CT images, however, are time consuming and less accurate due to the large amount of structural parameters. In this paper, a novel recognition method based on MobileNetV1 network, multiscale feature fusion method (MSFF), and dilated convolution is explored to deal with honeycomb lung in CT image classification. Firstly, the dilated convolution with different dilated rate is used to extract features to obtain receptive fields of different sizes, and then fuse the features of different scales at multiscale feature fusion block is used to solve the problem of feature loss and incomplete feature extraction. After that, by using linear activation functions (Sigmoid) instead of nonlinear activation functions (ReLu) in the improved deep separable convolution blocks to retain the feature information of each channel. Finally, by reducing the number of improved deep separable blocks to reduce the computation and resource consumption of the model. Results The experimental results show that improved MobileNet model has the best performance and the potential for recognition of honeycomb lung image datasets, which includes 6318 images. By comparing with 4 traditional models (SVM, RF, decision tree, and KNN) and 11 deep learning models (LeNet-5, AlexNet, VGG-16, GoogleNet, ResNet18, DenseNet121, SENet18, InceptionV3, InceptionV4, Xception, and MobileNetV1), our model achieved the performance with an accuracy of 99.52%, a sensitivity of 99.35%, and a specificity of 99.89%. Conclusion Improved MobileNet model is designed for the automatic recognition and classification of honeycomb lung in CT images. Through experiments comparative analysis of other models of machine learning and deep learning, it is proved that the proposed improved MobileNet method has the best recognition accuracy with fewer the model parameters and less the calculation time.

### 3. Cultural Tourism Industry Feature Extraction Based on Multiscale Feature

摘要: The cultural tourism industry combines the similarities of the cultural industry and tourism industry, which can be the most effective way to meet people's spiritual and cultural needs as well as their leisure needs simultaneously, and has a vast development potential. However, there are numerous and dispersed areas where the market value of cultural tourism resources is clustered, and frequently, each city has the clustering area with the highest market value concentration of cultural tourism resources. This feature of the spatial distribution of the market value of cultural tourism resources is significant for promoting the development of cultural tourism as a whole and constructing the industry's overall structure. It has broad application potential for extracting and differentiating cultural tourism industry characteristics. Texture feature extraction is typically performed using dual-tree complex wavelet transform (DT-CWT) and Gabor wavelet. In this paper, we propose a multiscale DT-CWT and Gabor-based method for identifying the cultural tourism industry. The method first decomposes the images of cultural tourism into multiscale space using a Gaussian pyramid, then extracts the multiscale features of the images using DT-CWT and Gabor, and lastly achieves feature fusion. Using a support vector machine (SVM) classifier to achieve classification, the effectiveness of the feature extraction method is determined. The experimental findings demonstrate that the method proposed in this paper can achieve a high rate of recognition.

### 4. Learning based multi-scale feature fusion for retinal blood vessels

摘要: Many eye-related diseases will lead to blindness or worse when it is lack of treatment in the early stages of the disease. Retinal vessel is important for doctors to detect eye diseases, even though the increase of some thin vessels may also mean the occurrence of certain diseases. Therefore, automatic retinal vessel segmentation is of great help to doctors in diagnosing diseases. In this paper, an automatic vessel segmentation method is proposed for retinal image, which is based on support vector machine combining multi-scale feature fusion model and B-COSFIRE filter response. Firstly, the inverted green channel image is enhanced by B-COSFIRE filter to strengthen bar-like vessel structures. Then the features are extracted by means of line operator in a multiresolution way, namely that each filtered image is down-sampled to cover a wider area, hence each sampled pixels can obtain not only the global but also local information. Then the final obtained features from three scales together along the depth direction are combined to train the SVM model. Finally, we use the classifier model to predict blood vessels. The proposed algorithm is evaluated on the public available fundus images datasets (DRIVE: Precision = 0.8657, Se = 0.7088, Sp = 0.9660 and ACC = 0.9900; STARE: Precision = 0.8782, Se = 0.6189, Sp = 0.9908 and ACC = 0.9494). The experiment results show that our proposed algorithm has effects on retinal vessels segmentation.  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1748-3018  
eISSN:  
1748-3026  
入藏号:  
WOS:000855671300001  
Indexed Date:  
2022-09-27  
Institution Address:  
Fujian Agr & Forestry Univ, Coll Comp & Sci, Fuzhou, Peoples R China  
Institution Address:  
Fujian Agr & Forestry, Coll Comp & Informat Sci, Inst Smart Agr & Forestry & Big Data, Fuzhou, Peoples R China  
Institution:  
Institution:

### 5. A novel multi-scale competitive network for fault diagnosis in rotating

摘要: Bearing fault diagnosis plays a vital role in ensuring the safe and reliable operation of rotating machinery. The diagnostic process is more difficult when the fault is in its early stages, as fewer fault components are contained in the vibration signal, making the diagnosis process more difficult. To improve the accuracy and efficiency of fault type identification, a novel multi-scale competitive network for fault diagnosis is proposed in this paper. First, to obtain multi-scale features and fully utilize the features in the intermediate layer, ensuring the completeness of fault information, a novel improved multi-scale feature fusion residual network (IMSFFRN) is proposed to exploit deep features for vibration signals. Specifically, multi-scale features are obtained by convolution with different dilation rates, and features from adjacent intermediate layers are selected for efficient fusion. Second, different features have varying importance in fault detection tasks. To make neurons more sensitive to specific faults, we propose a multiple-winning consciousness self-organizing map (MCSOM) competition layer, in which each neuron learns specific faults through competition, and the neuron that wins the competition updates its weights. This distinguishes the sensitivity of neurons to different faults. Finally, the generalizability of the network is improved by using support vector machines (SVMs) to classify fault classes. To affirm the efficacy of the proposed approach, a comprehensive evaluation is conducted on the CWRU, PU bearing dataset and SEU gear dataset. The results indicate that the accuracies achieved by the method proposed in this paper are 100%, 99.56% and 100%, respectively. It is superior to other methods proposed in this paper. Furthermore, it is observed that the proposed method exhibits high robustness in noisy environments.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
8  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
被引频次合计:  
8  
ISSN:  
0952-1976  
eISSN:  
1873-6769  
入藏号:  
WOS:001119054100001  
Indexed Date:  
2023-12-21  
Institution Address:  
Zhejiang Sci Tech Univ, Sch Informat Sci & Engn, Hangzhou 310018, Zhejiang, Peoples R China  
Institution:

### 6. Dynamic monitoring of land-use/land-cover change in cold and arid region

摘要: In this study, we conducted researchon the extraction of ground elements and the dynamic monitoring of land cover change in cold and arid areas with fragile ecological environments. The Mosuowan reclamation area in Xinjiang, China was selected as the target area, and the Landsat series satellite images from 2015 to 2019 were cut into subimages with sizes of 416\*416 px. Specifically, the training set, validation set and testset are 306 subimages, 204 subimages and 120 subimages respectively. Three traditional methods were then evaluated: MLC, SVM, and RF. Five semantic segmentation models were also evaluated: DeepLabv3+(Xception), DeepLabv3+ (MobileNet), SegNet (ResNet50), U-Net (MobileNet), and PSPNet (MobileNet). In the evaluation experiments, DeepLabv3+ (Xception) was found to achieve the optimal segmentation effect and multiscale feature fusion using fewer parameters. The overall accuracy, Kappa coefficient, precision, recall, F1-score, and MIoU were 96.06%, 0.96, 87.69%, 83.78%, 0.86, and 0.77, respectively. The MIoU of the DeepLabv3 + (Xception) model was significantly better than those of the other four models, improving from 0.03 to 0.39. On the basis of the land use classification results of the long-term timeseries remote sensing data from 1998 to 2020, we analyze spatial structure change in land use and the associated driving factors. Over the past 23 years, the total areas of desert, agricultural land, and construction land have been reduced by 15.00%, 12.68%, 2.53%, respectively. At the same time, the amount of water area has remained relatively stable. The overall transformation direction of land use is from desert to agricultural land and then from agricultural land to construction land. It can be seen from the results that the desertification control was effective, and urbanization rapidly developed. Consequently, this study can provide a reference for the application of deep learning in the field of medium-resolution remote sensing images, which can be used to realize the dynamic monitoring of land use and change.  
摘要:  
针对生态环境脆弱的寒旱区开展地物要素提取以及土地覆盖变化监测研究,对农业规划、城乡建设、生态环境监测与保护等具有重要意义。借助20152019年新疆莫索湾垦区Landsat-8影像构建数据集,对比3种传统方法:最大似然分类(Maximum likelihood classification,MLC)、支持向量机(Support vector machine,SVM)和随机森林(Random forest,RF)及5种语义分割模型:Deep-Labv3+(Xception)、DeepLabv3+(MobileNet)、SegNet(ResNet50)、U-Net(MobileNet)和PSPNet(MobileNet),选取最优自动化地物提取模型对研究区19982020年农用地、建筑用地、水体和荒漠4种地物要素进行分类,并运用土地利用转移矩阵和动态度进行定量动态变化分析。结果表明:Deep-Labv3+(Xception)模型可以实现更准确、更高效的地物提取,总体精确度(OA)、Kappa系数和F1值分别为96.06%、0.96和0.86,其中所选模型的平均交并比(MIoU)较其他模型提升0.03~0.39。近23 a,莫索湾垦区的荒漠、农用地和建筑用地三者的土地结构转化较为明显,荒漠总面积减少15.00%,农用地总面积增加12.68%,建筑用地总面积增加2.53%,水体面积变化较为平稳。地物类型总体转变方向为荒漠向农用地转化、农用地向建筑用地转化。该研究可为深度学习技术应用于中分辨率遥感卫星影像领域中实现土地利用及变化动态监测提供参考。  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
3  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
被引频次合计:  
3  
ISSN:  
1000-6060  
入藏号:  
CSCD:7114550  
Indexed Date:  
2022-02-25  
Institution Address:  
石河子大学信息科学与技术学院;;兵团空间信息工程技术研究中心;;兵团工业技术研究院, ;;兵团空间信息工程技术研究中心;;, 石河子;;石河子;;石河子, 新疆;;新疆;;新疆 832003;;832003;;832003, 中国  
Institution Address:  
安徽大学互联网学院, 合肥, 安徽 230039, 中国  
Institution Address:  
石河子大学信息科学与技术学院;;兵团空间信息工程技术研究中心, ;;兵团空间信息工程技术研究中心, 石河子;;石河子, 新疆;;新疆 832003;;832003, 中国  
Institution Address:  
School of Information Science and Technology, Shihezi University;;Research Center for Space Information Engineering Technology of XPCC;;XPCC Industrial Technology Research Institute, ;;Research Center for Space Information Engineering Technology of XPCC;;, Shihezi;;Shihezi;;Shihezi, Xinjiang;;Xinjiang;;Xinjiang 832003;;832003;;832003  
Institution Address:  
Internet College of Anhui University, Hefei, Anhui 230039, China  
Institution Address:  
School of Information Science and Technology, Shihezi University;;Research Center for Space Information Engineering Technology of XPCC, ;;Research Center for Space Information Engineering Technology of XPCC, Shihezi;;Shihezi, Xinjiang;;Xinjiang 832003;;832003  
Institution:  
石河子大学信息科学与技术学院;;兵团空间信息工程技术研究中心;;兵团工业技术研究院  
Institution:  
安徽大学互联网学院  
Institution:  
石河子大学信息科学与技术学院;;兵团空间信息工程技术研究中心  
Institution:  
School of Information Science and Technology, Shihezi University;;Research Center for Space Information Engineering Technology of XPCC;;XPCC Industrial Technology Research Institute  
Institution:  
Internet College of Anhui University  
Institution:  
School of Information Science and Technology, Shihezi University;;Research Center for Space Information Engineering Technology of XPCC

### 7. Chinese Food Recognition Model Based on Improved Residual Network

摘要: In view of the fact that traditional neural networks cannot effectively classify Chinese food with high similarity, a Chinese food recognition model of RNA-TL (RcsNct with attention and triplet loss) based on an improved residual network is proposed. The algorithm first fuses the multi-scale features to extract the semantic information of deep-level images, and then adds an attention mechanism layer to give more attention to the important parts of the images. Finally, the similarity among classes is calculated by using triplet-loss, whose result is input into support vector machine (SVM) for classification. The experimental results indicate that the proposed RNA-TI, model possesses more superior performances in recognition accuracy on the public dataset of Chinese food and the dataset collected by our project team, compared with the other mainstream algorithm models.  
在中国科学引文数据库中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
2  
ISSN:  
1006-4125  
入藏号:  
WOS:000686481200026  
Indexed Date:  
2021-09-03  
Institution Address:  
Nanjing Univ Informat Sci & Technol, Sch Automat, Nanjing 210011, Jiangsu, Peoples R China  
Institution:

## Autoencoder

### 1. A multi-channel ECG signal deep compressive sensing method using

摘要: Conventionally, multi-lead Electrocardiogram (ECG) signals are recorded and stored using high sensing rates and high precision, resulting in huge data volumes and placing greater pressure on storage and transmission resources. Compressive sensing (CS) allows efficient encoding and decoding of signals through sparse representation, measurement and reconstruction for compressed storage and transmission of ECG signals. Traditional CS methods generally requires manually selecting the features or dictionaries used in sparse representations, which lacks adaptivity and flexibility, and the complexity of the reconstruction process limits the application of CS in some scenarios with high real-time requirements. In this paper, we propose a deep compressive sensing framework for processing multi-lead ECG signals by combining CS and deep learning, which is based on multiscale feature fusion to construct a binary tree-shaped autoencoder architecture to achieve efficient compression and reconstruction of ECG signals. Experiments on the PTB diagnostic ECG database show that the proposed method achieves "very good" reconstruction quality at low sensing rates, with PRD and SNR of 1.77 % and 35.67 dB, respectively, when SR = 20 %. In addition, we investigate how the number of quantization bits affects the quality of reconstructed signals. After analysis, we found that 8-bit quantization can be used in practical applications to ensure the quality of reconstructed ECG signals, while decrease the bit rate of the data, improving the transmission efficiency, and reduce energy consumption.

### 2. SMD Anomaly Detection: A Self-Supervised Texture-Structure Anomaly

摘要: In electronic manufacturing, anomaly detection of surface mount devices (SMDs) through computer vision (CV) is an important task to control the production quality of SMDs. The difficulty of the detection is that some anomalous regions on the surfaces of SMDs are very minor and with variable shapes, which leads to poor detection efficiency. To solve this problem, based on the assumption that normal samples can be reconstructed more accurately than anomalous samples, a self-supervised image anomaly detection framework with a multiscale two-branch feature fusion strategy is proposed. Specifically, it adopts autoencoder (AE) as the basic framework, and to enhance the reconstruction error between input anomalous samples and the reconstructed ones, a self-supervised learning task of reconstructing images is introduced to have the model neglect the encoding of the suspected anomalous regions found by a contextual attention mask (CAM) module. Meanwhile, a multiscale feature fusion strategy is developed to fuse texture and structure features in the decoder to reconstruct samples. Moreover, a multilevel anomalous score criterion is proposed to enlarge the scores for the samples with very minor anomalies. At last, an SMD-capacitor anomaly detection dataset (SMDC-DET) is built to evaluate the proposed method. The experiments show that the proposed method achieves an average area under the curve (AUC) accuracy of 98.82%, much better when compared to the start-of-the-art existing anomaly detection methods.

### 3. IoT-based prediction model for aquaponic fish pond water quality using

摘要: The Internet of Things (IoT)-based smart solutions have been developed to predict water quality and they are becoming an increasingly important means of providing efficient solutions through communication technologies. IoT systems are used for enabling connection between various devices based on the ability to gather and collect information. Furthermore, IoT systems are designed to address the environment and the automation industry. The threats associated with aquaponics farming are managed through an IoT-based smart water monitoring framework, which has become increasingly relevant in recent days. Therefore, this approach is crucial for achieving a remarkable improvement in order to increase the productivity rate and yield. The quality of water directly affects the rate of growth, efficiency of feed, and the overall health rate of the fish, plants, and bacteria. Insufficient knowledge about species selection poses a significant challenge in aquaponics farming, as it heavily relies on the water quality parameters. To address the challenges of conventional models, we have developed an effective IoT-based water quality prediction model, more specifically designed for aquaponic fish ponds. The data needed to perform the developed water quality prediction model will be acquired from "a simple dataset of aquaponic fish pond IoT" database. After that, these data are forwarded to the feature extraction phase. The weighted features, DBN (Deep Belief Network) features, and the original features are achieved in the feature extraction stage. The weighted features are obtained using the Revamped Fitness-based Mother Optimization Algorithm (RF-MOA). Subsequently, these extracted features are fed into the Multi-Scale feature fusion-based Convolutional Autoencoder with a Gated Recurrent Unit (MS-CAGRU) network for predicting the water quality. Thus, the water quality predicted data is obtained. The proposed model integrates GRU networks with a convolutional autoencoder to improve water quality prediction by capturing trends and managing temporal dependencies. It enhances accuracy by analysing key parameters and employing techniques to reduce overfitting. The effectiveness of the proposed system is evaluated in comparison to the traditional models using some evaluation measures.

### 4. MaskRecon: High-quality human reconstruction via masked autoencoders

摘要: In this paper, we explore reconstructing high-quality clothed 3D humans from a single RGB-D image, assuming that virtual humans can be represented by front-view and back-view depths. Due to the scarcity of captured real RGB-D human images, we employ rendered images to train our method. However, rendered images lack background with significant depth variation in silhouettes, leading to shape prediction inaccuracies and noise. To mitigate this issue, we introduce a pseudo-multi-task framework, which incorporates a Conditional Generative Adversarial Network (CGAN) to infer back-view RGB-D images and a self-supervised Masked Autoencoder (MAE) to capture latent structural information of the human body. Additionally, we propose a Multi-scale Feature Fusion (MFF) module to effectively merge structural information and conditional features at various scales. Our method surpasses many existing techniques, as demonstrated through evaluations on the Thuman, RenderPeople, and BUFF datasets. Notably, our approach excels in reconstructing high-quality human models, even under challenging conditions such as complex poses and loose clothing, both on rendered and real-world images. Codes are available at https://github.com/Archaic-Atom/MaskRecon.  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
0925-2312  
eISSN:  
1872-8286  
入藏号:  
WOS:001307872400001  
Indexed Date:  
2024-09-14  
Institution Address:  
Northwestern Polytech Univ, Sch Elect & Informat, Xian 710129, Peoples R China  
Institution Address:  
Nanchang Hangkong Univ, Sch Informat Engn, Nanchang 330063, Peoples R China  
Institution Address:  
Northwestern Polytech Univ, Sch Comp Sci, Xian 710129, Peoples R China  
Institution Address:  
Content Prod Ctr Virtual Real, Beijing 101318, Peoples R China  
Institution:  
Institution:  
Institution:  
Institution:  
Content Prod Ctr Virtual Real

### 5. Joint blind image deblurring and super-resolution via double-branch

摘要: Single image super-resolution (SISR) has recently received considerable attention and has made great progress. The goal of SISR is to generate a high-resolution (HR) image from a low-resolution (LR) image. Many visual tasks, such as object detection,1 image dehazing,2 image deraining,3,4 and facial recognition,5 require HR images to improve their recognition accuracy. However, image degradation, such as a blurred image, is inevitable due to the shaking of the camera or movement of the target. Degradation images will severely reduce the recognition accuracy of target detection and other tasks. In addition, the variation of moving target boundary leads to complex blur information in the image. Therefore, images captured from real scenes are negatively affected by nonuniform blur during acquisition and processing. Blurred images with unknown blur kernels present a great challenge for the super-resolution (SR) task and many other high-level visual tasks.  
Unexpected situations, such as object movement and camera shaking, cause interference and produce blurry, degraded images. These degraded images can adversely affect visual processing tasks, such as target detection and facial recognition. In addition, the super-resolution (SR) of a single image is a basic task to improve the image quality in visual applications. However, due to the existence of artifacts, it is difficult to generate clear SR images from rough and degraded images. In this work, we propose a double-branch projection feedback network that can generate clear high-resolution (HR) images from severely blurred low-resolution (LR) images. One is a deep feature projection feedback branch that can obtain texture-rich SR feature maps and reduce the accumulative errors with depth, and a channel adjustment and multi-scale feature fusion mechanism are proposed to better obtain the mapping relation between LR and HR images. The other is an autoencoder deblurring module, which is used to get deblurred features of an input image. Moreover, a mapping associated feature block is applied to match non-local feature information in the feature fusion module. Extensive experiments show that the network is effective and performs better than the state-of-the-art methods in single image deblurring and SR. (c) 2021 SPIE and IS&T [DOI: 10.1117/1.JEI.30.2.023033]  
SciELO Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
0  
ISSN:  
1017-9909  
eISSN:  
1560-229X  
入藏号:  
WOS:000655241000002  
Indexed Date:  
2021-06-04  
Institution Address:  
Shanghai Univ Engn Sci, Sch Elect & Elect Engn, Control Engn, Shanghai, Peoples R China  
Institution Address:  
Shanghai Comp Software Technol Dev Ctr, Shanghai, Peoples R China  
Institution:  
Institution:  
Shanghai Comp Software Technol Dev Ctr

### 6. A multi-scale dual-decoder autoencoder model for domain-shift machine

摘要: Anomaly detection through machine sounds plays a crucial role in the development of industrial automation due to its excellent flexibility and real-time response capabilities. However, in real-world scenarios, the occurrence frequency of machine anomaly events is relatively low, making it difficult to collect anomaly sound data under various operating conditions. Moreover, due to the influence of operating conditions and environmental noise, the collected sound data may have distribution differences, leading to data domain shifts issues. To address these problems, we propose an unsupervised multi-scale dual-decoder autoencoder (MS-D2AE) network for anomaly sound detection. The MS-D2AE model consists of residual layers, an encoder, and two decoders. The model fuses fine-grained information of sound features through the Multi-scale Feature Fusion Module (MTSFFM), enabling the model to effectively learn feature data from multiple scales. By using a residual layer composed of a single MTSFFM, the encoder's input is directly connected to the intermediate results, further enhancing information transmission. The designed dual-decoder autoencoder structure, in addition to reconstructing error calculation, also utilizes the similarity error calculation between the outputs of the two decoders, encouraging the model to more accurately reconstruct the feature data during learning, thus more comprehensively learning the feature representation of normal data. Additionally, to mitigate the impact of data shift on model performance, we design a feature domain mixing method that blends sound features from both source and target domains to enhance the diversity and generalization of sound features. Finally, we have verified the effectiveness of this method on the Dcase2023 Challenge Task2 and Dcase2022 Challenge Task2 datasets.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
0  
ISSN:  
1051-2004  
eISSN:  
1095-4333  
入藏号:  
WOS:001336408800001  
Indexed Date:  
2024-10-27  
Institution Address:  
Hefei Univ, Dept Artificial Intelligence & Big Data, Hefei 230601, Anhui, Peoples R China  
Institution:

### 7. Probabilistic autoencoder with multi-scale feature extraction for

摘要: Effectively detecting anomalies for multivariate time series is of great importance for the modern industrial system. Recently, reconstruction-based deep learning methods have been widely used in time series anomaly detection. However, the rich local and global characteristics of time series may not be well captured by methods that compress and reconstruct time series through a single-scale neural network. In addition, under the influence of a complex environment, small fluctuations occur during the normal operation of the system, which will also bring challenges for those methods to reconstruct exact values. In this paper, we propose an unsupervised multivariate time series anomaly detection method based on a probabilistic autoencoder with multi-scale feature extraction (PAMFE). The multiple parallel dilated convolutions with different dilation factors and feature fusion module enable PAMFE to capture overall and detailed information of time series to identify various types of anomalies better. Furthermore, considering the normal fluctuation of data, we reconstruct the expected distribution of input and calculate the anomaly score based on the probability that the input belongs to the distribution. Extensive experiments on four publicly real-world datasets demonstrate that PAMFE outperforms state-of-the-art methods in F1-Score. Moreover, we investigate the contributions of the major components of PAMFE, and the experimental results show that they all contribute to performance improvement.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
1  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
3  
被引频次合计:  
4  
ISSN:  
0924-669X  
eISSN:  
1573-7497  
入藏号:  
WOS:000890125400002  
Indexed Date:  
2022-12-14  
Institution Address:  
Beijing Univ Posts & Telecommun, Sch Artificial Intelligence, Beijing 100876, Peoples R China  
Institution Address:  
Prod & Operat Dept, PipeChina West East Gas Pipeline Co, Shanghai 200122, Peoples R China  
Institution:  
Institution:  
Prod & Operat Dept

### 8. A target spatial location method for fuze detonation point based on deep

摘要: The spatial location of fuze detonation point is crucial for evaluating the working condition and improving the performance of fuze. Considering the observation safety, the non-contact long-distance accurate measurement technology is essential. In this paper, we propose a method that takes the sensor data of optics, spatial attitude and GPS as input and outputs the spatial position of the fuze detonation point. The proposed method consists of two steps. First, an object detection algorithm with post-processing algorithm is proposed to obtain rich information of the target. The algorithm achieves high-accuracy detection by introducing powerful backbone, attention mechanism, group convolution, and improved multi-scale feature fusion. Second, a Variational AutoEncoder (VAE) algorithm model improved by dense connection structure and multiple source heterogeneous sensor information fusion structure is proposed as the position regression algorithm. It receives the status information of the observer camera and the output of the object detection algorithm, and then outputs the threedimensional coordinates of the explosion point. Finally, method validation and performance analysis are realized through virtual scene simulation. Experiment results show the superiority of the proposed object detection algorithm over other typical algorithms on explosion flare detection, with its Average Precision (AP) of 0.889. The positioning error of the spatial location method is 0.896 m, while that of the binocular stereo vision method is 2.863 m. Therefore, the proposed target spatial location method is proved to be effective and accurate.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
4  
ISSN:  
0957-4174  
eISSN:  
1873-6793  
入藏号:  
WOS:001106812500001  
Indexed Date:  
2024-02-23  
Institution Address:  
Beijing Inst Technol, Sch Elect & Mech, Beijing 100081, Peoples R China  
Institution Address:  
Sci & Technol Elect Dynam Control Lab, Beijing, Peoples R China  
Institution Address:  
Beijing Inst Technol, Tangshan Res Inst, Tangshan 063611, Hebei, Peoples R China  
Institution:  
Institution:  
Sci & Technol Elect Dynam Control Lab  
Institution:

### 9. Image Multi-Scale Feature Compression Algorithm for Machine Vision Tasks

摘要: For the problems that,in the collaborative intelligence framework,the intermediate feature data of machine vision tasks is large and difficult to transmit efficiently,a multi-scale image feature fusion compression algorithm was proposed.Firstly,a cascaded residual transformation module was designed according to the multi-scale features output by the deep learning model on the edge device,the redundancy of multi-scale features was eliminated by stepwise subtraction of features of different sizes,and the residual features were compressed to a unified size.Then,an autoencoder was designed to eliminate the statistical redundancy of compact features by arithmetic coding.Next,a prediction and reconstruction module was designed on the cloud according to the compact features of decoding to generate the prediction features,which were combined with the residual features to accurately reconstruct the multi-scale features.Finally,a joint optimization function was built for the collaborative optimization of the modules including residual transformation,autoencoder,and prediction reconstruction,thus achieving the optimal trade-off between transmission bit rate and information representation ability.The simulation results show that the proposed algorithm has not only the largest feature compression ratio,but also the most complete reconstructed features in the space compression,and that when the transmission bit rate is 0.lbpp,the model accuracy of the proposed algorithm is improved by 8.57% and 3.87%,respectively,compared with the image coding algorithm VVC and the feature compression algorithm MSFC.This study can provide technical support for the coding framework of machine vision,and has certain value in engineering application.

## Principal Component Analysis (PCA)

### 1. Edge-Aware Multiscale Feature Integration Network for Salient Object

摘要: The optical remote sensing images (RSIs) show various spatial resolutions and cluttered background, where salient objects with different scales, types, and orientations are presented in diverse RSI scenes. Therefore, it is inappropriate to directly extend cutting-edge saliency detection methods for conventional RGB images to optical RSIs. Besides, the existing saliency models targeting RSIs often render imperfect saliency maps, where some of them are with coarse boundary details. To solve this problem, this article attempts to introduce the edge information to precisely detect salient objects in RSIs. Accordingly, we propose an edge-aware multiscale feature integration network (EMFI-Net) for salient object detection by conducting multiscale feature integration under the explicit and implicit assistance of salient edge cues. Specifically, our network contains two parts including the encoder and decoder. First, the encoder extracts multiscale deep features from three RSIs with different resolutions, where the high-level deep semantic features from three RSIs are integrated using a cascaded feature fusion module. Second, the encoder explicitly enriches the multiscale deep features by integrating the salient edge cues extracted by a salient edge extraction module. Meanwhile, we also implicitly deploy an edge-aware constraint to the supervision of the saliency map prediction by introducing a hybrid loss function. Finally, the decoder integrates the enriched multiscale deep features in a coarse-to-fine way, yielding a high-quality saliency map. The experiments conducted on two public optical RSI datasets clearly prove the effectiveness and superiority of the proposed EMFI-Net against the state-of-the-art saliency models.

### 2. Enhancing multi-scale information exchange and feature fusion for human

摘要: Multi-scale feature fusion is an important part of modern network architectures to extract more comprehensive information for most computer vision tasks, such as semantic segmentation and keypoint estimation. However, most existing multi-scale methods add fusion connections between layers or branches directly, which inevitably ignores the semantic information discrepancy between feature maps with different resolutions and depths. Moreover, inappropriate fusion connections may lead to the loss of channel-wise and spatial information. In this paper, we propose a method to enhance and refine multi-scale feature fusion for human pose estimation by employing two attention mechanisms. Specifically, we present a novel multi-head spatial attention (MHSA), which is employed to model context information of the intermediate feature maps and reinforce important local features. Meanwhile, we utilize the position channel attention (PCA) to capture long-range dependencies while retaining the important position information in the attention maps. Combining with the modules of MHSA and PCA, we design an enhanced multi-scale feature fusion network (EMF-HRNet) based on the high-resolution network (HRNet). Our proposed EMF-HRNet yields better results with repeated multi-scale information exchange and feature fusion units. Extensive experiments on two common benchmarks, COCO Keypoint dataset and MPII Human Pose dataset, show that our method significantly improves the performance of state-of-the-art pose estimation methods.

### 3. Fine-grained image recognition based on mid-level subtle feature

摘要: In the field of fine-grained visual recognition, due to subtle differences between highly similar categories, precise extraction of subtle image features has a crucial impact on recognition accuracy. It has become a trend for the existing related hot research algorithms to use attention mechanism to extract categorical features, however, these algorithms ignore the subtle but distinguishable features, and isolate the feature relationships between different discriminative regions of objects. Aiming at these problems, a fine-grained image recognition algorithm based on mid-level subtle feature extraction and multi-scale feature fusion was proposed. First, the salient features of image were extracted by using the weight variance measures of channel and position information fused mid-level features. Then, the mask matrix was obtained through the channel average pooling to suppress salient features and enhance the extraction of subtle features in other discriminative regions. Finally, channel weight information and pixel complementary information were used to obtain multi-scale fusion features of channels and pixels to enhance the diversity and richness of different discriminative regional features. Experimental results show that the proposed algorithm achieves 89.52% Top-1 accuracy and 98.46% Top-5 accuracy on dataset CUB-200-211, and 94.64% Top-1 accuracy and 98.62% Top-5 accuracy on dataset Stanford Cars, and 93.20% Top-1 accuracy and 97.98% Top-5 accuracy on dataset Fine-Grained Visual Classification of Aircraft (FGVC-Aircraft). Compared with recurrent collaborative attention feature learning network PCA-Net (Progressive Co-Attention Network) algorithm, the proposed algorithm has the Top-1 accuracy increased by 1.22, 0.34 and 0.80 percentage points respectively, and the Top-5 accuracy increased by 1.03, 0.88 and 1.12 percentage points respectively.

### 4. Feature Aggregation and Propagation Network for Camouflaged Object

摘要: Camouflaged object detection (COD) aims to detect/segment camouflaged objects embedded in the environment, which has attracted increasing attention over the past decades. Although several COD methods have been developed, they still suffer from unsatisfactory performance due to the intrinsic similarities between the foreground objects and background surroundings. In this paper, we propose a novel Feature Aggregation and Propagation Network (FAP-Net) for camouflaged object detection. Specifically, we propose a Boundary Guidance Module (BGM) to explicitly model the boundary characteristic, which can provide boundary-enhanced features to boost the COD performance. To capture the scale variations of the camouflaged objects, we propose a Multi-scale Feature Aggregation Module (MFAM) to characterize the multi-scale information from each layer and obtain the aggregated feature representations. Furthermore, we propose a Cross-level Fusion and Propagation Module (CFPM). In the CFPM, the feature fusion part can effectively integrate the features from adjacent layers to exploit the cross-level correlations, and the feature propagation part can transmit valuable context information from the encoder to the decoder network via a gate unit. Finally, we formulate a unified and end-to-end trainable framework where cross-level features can be effectively fused and propagated for capturing rich context information. Extensive experiments on three benchmark camouflaged datasets demonstrate that our FAP-Net outperforms other state-of-the-art COD models. Moreover, our model can be extended to the polyp segmentation task, and the comparison results further validate the effectiveness of the proposed model in segmenting polyps. The source code and results will be released at https://github.com/taozh2017/FAPNet.  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
7  
Web of Science 核心合集中的 "被引频次":  
106  
在中国科学引文数据库中的被引频次:  
9  
被引频次合计:  
112  
ISSN:  
1057-7149  
eISSN:  
1941-0042  
入藏号:  
WOS:000888975000005  
PubMed ID:  
36331642  
Indexed Date:  
2022-12-04  
Institution Address:  
Nanjing Univ Sci & Technol, Key Lab Intelligent Percept & Syst High Dimens In, PCA Lab, Minist Educ, Nanjing 210094, Peoples R China  
Institution Address:  
Nanjing Univ Sci & Technol, Sch Comp Sci & Engn, Jiangsu Key Lab Image & Video Understanding Social, Nanjing 210094, Peoples R China  
Institution Address:  
Minist Educ, Key Lab Syst Control & Informat Proc, Shanghai 200240, Peoples R China  
Institution Address:  
Southeast Univ, Sch Comp Sci & Engn, Nanjing 211189, Peoples R China  
Institution Address:  
Lehigh Univ, Dept Bioengn, Bethlehem 18015, PA USA  
Institution:  
Institution:  
Institution:  
Minist Educ  
Institution:  
Institution:

### 5. ArcMask: a robust and fast image-based method for high-speed railway

摘要: The pantograph-catenary arcing reflects the health of pantograph-catenary and current collection quality of high-speed railway, so the arc detection is of great significance. However, due to the scene complexity, intra-class polymorphism and inter-class similarity of arcing and the fast running speed of high-speed railway, it is still a huge challenge to achieve fine and robust arcing detection. To overcome these issues, a robust and fast image-based instance segmentation method called ArcMask is proposed to detect pantograph-catenary arcing, which designs a new attention-based multi-scale feature fusion module that combines both top-down and down-up modules to realize arcing pixel-level instance segmentation. The effective combination of instance-level information and bottom-level semantic information balances features representation ability of top-level and bottom-level features. Compared with other instance segmentation methods (e.g., BlendMask), it can effectively learn feature representation with tiny, irregular and complex arc features and speeds up the calculation. In addition, both deformable convolution and depth-wise separable convolution are introduced in ArcMask, which aims to improve the segmentation performance of irregular arcing and efficiency. The ArcMask can distinguish different arcing instances at pixel-level with fine granularity and distinguish inter-class and intra-class features of arcing, instead of just focusing on rectangular bounding box. Experiments on self-collected dataset IVAIS-PCA2021 verify the effectiveness and efficiencies of the ArcMask. Its AP, AP(50) and AP(75) are 56.61, 94.14 and 64.56, respectively, and the fastest reasoning speed based on MobileNet is 56 FPS. Compared with other state-of-the-art segmentation methods, the proposed ArcMask has better integrity in arcing edge detection.  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
4  
被引频次合计:  
4  
ISSN:  
0941-0643  
eISSN:  
1433-3058  
入藏号:  
WOS:000913290300001  
Indexed Date:  
2023-02-01  
Institution Address:  
Southwest Jiaotong Univ, Engn Technol Res Ctr, Natl Rail Transit Electrificat & Automation, Chengdu, Sichuan, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, Sch Elect Engn, Chengdu, Sichuan, Peoples R China  
Institution Address:  
Southwest Jiaotong Univ, State Key Lab Tract Power, Chengdu, Sichuan, Peoples R China  
Institution:  
Institution:  
Institution:

## Graph Neural Network (GNN)

### 1. Multivariate Time Series Forecasting Based on Multi-Scale Feature Fusion

摘要: Each subsequence of the Multivariate Time Series(MTS) contains multi-scale characteristics of different time spans,comprising information such as development process,direction,and trend. However,existing time series prediction models cannot effectively capture multi-scale features and evaluate their importance. In this study,a MTS prediction network,FFANet,is proposed based on multi-scale temporal feature fusion and a Dual-Attention Mechanism (DAM).FFANet effectively integrates multi-scale features and focuses on important parts.Utilizing the parallel temporal dilation convolution layer in the multi-scale temporal feature fusion module endows the model with multiple receptive domains to extract features of temporal data at different scales and adaptively fuse them based on their importance. Using a DAM to recalibrate the fused temporal features,FFANet focuses on features that make significant contributions to prediction by assigning temporal and channel attention weights and weighting them to the corresponding temporal features. The experimental results show that compared with AR,VARMLP,RNN-GRU,LSTNet-skip,TPA-LSTM, MTGNN,and AttnAR time series prediction models,FFANet achieves average reduction of 0.152 3、0.120 0、0.074 3、 0.035 4、0.021 5、0.012 1、0.020 0 in RRSE prediction error on Traffic,Solar Energy,and Electricity datasets,respectively.  
摘要:  
多元时间序列的各子序列包含不同时间跨度的多尺度特征,现有时间序列预测模型不能有效地捕获多尺度特征以及评估其重要程度。提出一种基于多尺度时序特征融合与双注意力机制的多元时间序列预测网络FFANet,有效融合多尺度特征并关注其中重要部分。通过多尺度时序特征融合模块中并行的时序膨胀卷积层,使模型具有多种感受域,从而提取时序数据在不同尺度上的特征,并根据重要性对其进行自适应融合。利用双注意力模块对融合的时序特征进行重新标定,通过分配时序和通道注意力权重并加权至对应的时序特征,使FFANet聚焦对预测有重要贡献的特征。实验结果表明,相比AR、VARMLP、RNN-GRU、LSTNet-skip、TPA-LSTM、MTGNN和AttnAR时间序列预测模型,FFANet在Traffic、Solar Energy和Electricity数据集上的RRSE预测误差分别平均降低0.152 3、0.120 0、0.074 3、0.035 4、0.021 5、0.012 1、0.020 0。  
在 BIOSIS Citation Index 中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
1000-3428  
入藏号:  
CSCD:7569525  
Indexed Date:  
2023-12-29  
Institution Address:  
中国民航大学计算机科学与技术学院, 天津 300300, 中国  
Institution Address:  
潍坊学院计算机工程学院, 潍坊, 山东 261061, 中国  
Institution Address:  
School of Computer Science and Technology,Civil Aviation University of China, Tianjin 300300, China  
Institution Address:  
School of Computer Engineering,Weifang University, Weifang, Shandong 261061, China  
Institution:  
中国民航大学计算机科学与技术学院  
Institution:  
潍坊学院计算机工程学院  
Institution:  
School of Computer Science and Technology,Civil Aviation University of China  
Institution:  
School of Computer Engineering,Weifang University

### 2. Vision graph convolutional network for underwater image enhancement

摘要: Colour deviation, non -uniform degradation, and decreased contrast often occur in underwater images because a certain amount of light is absorbed and dispersed underwater. To address this problem, a graph convolutionbased underwater image enhancement method (GC-UIE) is proposed. Specifically, patches of underwater images are treated as graph structure, and low -quality underwater images are enhanced by leveraging the advantages of vision graph neural network (VIG). Considering the distortion of underwater images in detail and colour, a local multi -scale feature fusion module and a colour channel correction module based on the mechanism of self -attention are proposed and embedded into the network. Furthermore, the local features are extracted using a convolutional model with multiple receptive fields to complement the global features. To improve colour quality, a self -attention mechanism is utilized. Finally, the underwater images are restored using a residual connection design based on the underwater imaging models. The GC-UIE performed, both qualitatively and quantitatively, better than the other methods. The PyTorch code will be available at https://github.com/xzx11/GC-UIE.  
Russian Science Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
2  
Arabic Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
2  
ISSN:  
0950-7051  
eISSN:  
1872-7409  
入藏号:  
WOS:001254396300001  
Indexed Date:  
2024-07-02  
Institution Address:  
Ningbo Univ, Sch Math & Stat, Ningbo 315211, Peoples R China  
Institution Address:  
Ningbo Univ, Fac Informat Sci & Engn, Ningbo 315211, Peoples R China  
Institution:  
Institution:

### 3. Graph clustering network with structure embedding enhanced

摘要: Recently, deep clustering utilizing Graph Neural Networks has shown good performance in the graph clustering. However, the structure information of graph was underused in existing deep clustering methods. Particularly, the lack of concern on mining different types structure information simultaneously. To tackle with the problem, this paper proposes a Graph Clustering Network with Structure Embedding Enhanced (GC-SEE) which extracts nodes importance-based and attributes importance-based structure information via a feature attention fusion graph convolution module and a graph attention encoder module respectively. Additionally, it captures different orders-based structure information through multi-scale feature fusion. Finally, a self -supervised learning module has been designed to integrate different types structure information and guide the updates of the GC-SEE. The comprehensive experiments on benchmark datasets commonly used demonstrate the superiority of the GC-SEE. The results showcase the effectiveness of the GC-SEE in exploiting multiple types of structure for deep clustering.  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
2  
Web of Science 核心合集中的 "被引频次":  
17  
在 BIOSIS Citation Index 中的被引频次:  
0  
SciELO Citation Index 中的 "被引频次":  
0  
被引频次合计:  
19  
ISSN:  
0031-3203  
eISSN:  
1873-5142  
入藏号:  
WOS:001057454200001  
Indexed Date:  
2023-09-12  
Institution Address:  
China Univ Min & Technol, Sch Comp Sci & Technol, Xuzhou 221116, Peoples R China  
Institution Address:  
Minist Educ Peoples Republ China, Mine Digitizat Engn Res Ctr, Xuzhou 221116, Peoples R China  
Institution Address:  
Tianjin Univ, Coll Intelligence & Comp, Tianjin 300350, Peoples R China  
Institution:  
Institution:  
Minist Educ Peoples Republ China  
Institution:

### 4. MSGAFN: Multi-scale graph attention fusion network for machine fault

摘要: Fault diagnosis is essential in various fields, such as industrial manufacturing and engineering maintenance. Graph neural networks that take graph data as input can explore the relationships between data, which has strong feature expression capabilities. Currently, traditional fault diagnostic methods based on graph neural networks (GNN) are challenging to capture local and global feature information of data effectively. Most GNN models fail to consider the inherent differences between adjacent nodes, and they perform poorly in processing vibration signals in real-world industrial scenarios that commonly have strong noise. To resolve these concerns, this paper proposes a method for fault diagnosis based on a multi-scale graph attention fusion network (MSGAFN). In MSGAFN, data samples are constructed as clan graphs with multiple information scales to effectively represent local and global information of the graph structure data. Additionally, MSGAFN designed a new multi-scale feature fusion layer (MSFFL) to automatically learn the weights of adjacent nodes to represent their importance to the central node and reflect the differences between different adjacent nodes. This method is fully validated on bearing and gear datasets. The experimental results demonstrate that the proposed method exhibits excellent performance under conditions of imbalanced datasets and strong noise, providing a promising approach for bearings and gears fault diagnosis in real-world industrial scenarios with strong noise.  
Arabic Citation Index 中的 "被引频次":  
0  
在中国科学引文数据库中的被引频次:  
0  
Web of Science 核心合集中的 "被引频次":  
1  
SciELO Citation Index 中的 "被引频次":  
0  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
被引频次合计:  
1  
ISSN:  
0954-4062  
eISSN:  
2041-2983  
入藏号:  
WOS:001174651600001  
Indexed Date:  
2024-03-09  
Institution Address:  
Qingdao Univ Technol, Sch Informat & Control Engn, Qingdao, Shandong, Peoples R China  
Institution Address:  
Qingdao Univ Technol, Sch Informat & Control Engn, 777 Jialingjiang East Rd, Qingdao 266033, Shandong, Peoples R China  
Institution:  
Institution:

### 5. GMAlignNet: multi-scale lightweight brain tumor image segmentation with

摘要: Although the U-shaped architecture, represented by UNet, has become a major network model for brain tumor segmentation, the repeated convolution and sampling operations can easily lead to the loss of crucial information. Additionally, directly fusing features from different levels without distinction can easily result in feature misalignment, affecting segmentation accuracy. On the other hand, traditional convolutional blocks used for feature extraction cannot capture the abundant multi-scale information present in brain tumor images. This paper proposes a multi-scale feature-aligned segmentation model called GMAlignNet that fully utilizes Ghost convolution to solve these problems. Ghost hierarchical decoupled fusion unit and Ghost hierarchical decoupled unit are used instead of standard convolutions in the encoding and decoding paths. This transformation replaces the holistic learning of volume structures by traditional convolutional blocks with multi-level learning on a specific view, facilitating the acquisition of abundant multi-scale contextual information through low-cost operations. Furthermore, a feature alignment unit is proposed that can utilize semantic information flow to guide the recovery of upsampled features. It performs pixel-level semantic information correction on misaligned features due to feature fusion. The proposed method is also employed to optimize three classic networks, namely DMFNet, HDCNet, and 3D UNet, demonstrating its effectiveness in automatic brain tumor segmentation. The proposed network model was applied to the BraTS 2018 dataset, and the results indicate that the proposed GMAlignNet achieved Dice coefficients of 81.65%, 90.07%, and 85.16% for enhancing tumor, whole tumor, and tumor core segmentation, respectively. Moreover, with only 0.29 M parameters and 26.88G FLOPs, it demonstrates better potential in terms of computational efficiency and possesses the advantages of lightweight. Extensive experiments on the BraTS 2018, BraTS 2019, and BraTS 2020 datasets suggest that the proposed model exhibits better potential in handling edge details and contour recognition.  
在 BIOSIS Citation Index 中的被引频次:  
0  
Russian Science Citation Index 中的 "被引频次":  
0  
Arabic Citation Index 中的 "被引频次":  
0  
SciELO Citation Index 中的 "被引频次":  
0  
Web of Science 核心合集中的 "被引频次":  
1  
在中国科学引文数据库中的被引频次:  
0  
被引频次合计:  
1  
ISSN:  
0031-9155  
eISSN:  
1361-6560  
入藏号:  
WOS:001229818000001  
PubMed ID:  
38657628  
Indexed Date:  
2024-05-31  
Institution Address:  
Inner Mongolia Univ Sci & Technol, Sch Digital & Intelligent Ind, Inner Mongolia Key Lab Pattern Recognit & Intellig, Baotou, Peoples R China  
Institution Address:  
Inner Mongolia Univ Technol, Sch Informat Engn, Hohhot 010051, Peoples R China  
Institution:  
Institution: