***Image segmentation using U-Net and Mask R-CNN architectures***

*Akashchidambar Anilkumar Sangal*

Dept of Electronics and Communication Engineering

KLS Gogte Institute Of Technology

Belagavi,Karnataka India

akashsangal07@gmail.com

***Abstract:In this study, U-Net and Mask R-CNN, two well-liked architectures for image segmentation tasks, are compared and contrasted. Partitioning an image into several parts is a fundamental operation in computer vision known as image segmentation. In this field, U-Net and Mask R-CNN have been extensively utilised and successfully shown. The purpose of this study is to assess each method's applicability for various picture segmentation applications by comparing their performance, accuracy, and efficiency on diverse datasets. We give insights into the advantages and disadvantages of both designs by thorough testing and assessment, empowering academics and practitioners to choose between U-Net and Mask R-CNN for particular image segmentation applications.***

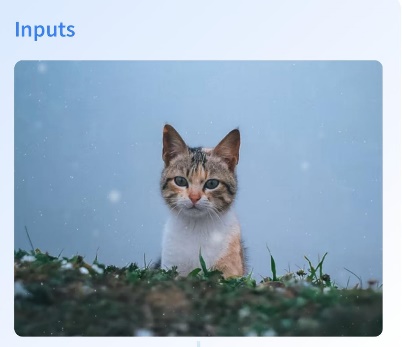
I. INTRODUCTION

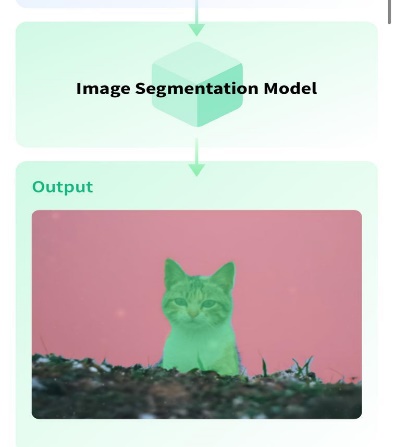
Partitioning a picture into many coherent sections, or "segments," is a crucial operation in computer vision that enables the accurate detection and extraction of objects or areas of interest. For many applications, such as object detection, medical imaging, autonomous driving, and augmented reality, precise picture segmentation is crucial.

By offering incredibly effective and economical solutions, deep learning algorithms have recently revolutionised the area of picture segmentation. U-Net and Mask R-CNN are two well-known deep learning architectures that are frequently used for image segmentation.

1.1 Image segmentation :

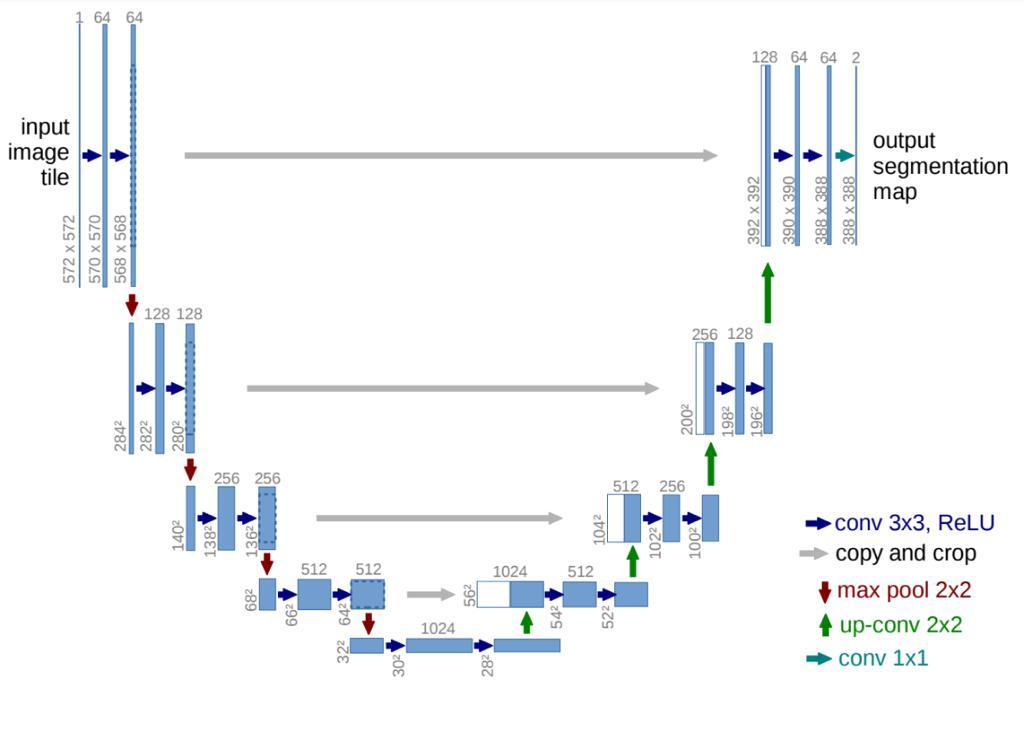
Image Segmentation divides an image into segments where each pixel in the image is mapped to an object. This task has multiple variants such as instance segmentation, panoptic segmentation and semantic segmentation.





1.2 U -Net:

It was first suggested to use U-Net, a fully convolutional neural network architecture, for segmenting biological images. Due to its outstanding performance, it has now grown in popularity across several sectors. An encoder-decoder structure with skip links is a defining feature of the U-Net architecture. While the decoder section recovers the spatial information and creates the segmentation map, the encoder part extracts the context and abstract features from the input picture. Skip connections make it possible for high-resolution characteristics to be sent directly from the encoder to the associated decoder layers, allowing accurate localization and the retention of fine-grained information. U-Net is renowned for its ability to deliver precise segmentation results and has demonstrated exceptional usefulness in situations with little training data.



1.3 Mask R-CNN:

The Faster R-CNN object identification framework has an addition called Mask R-CNN that includes instance-level segmentation. Along with object detection, Mask R-CNN also creates pixel-level masks for each item that is discovered, offering thorough segmentation data. This is accomplished by merging a fully convolutional network for mask creation with a region proposal network (RPN) for object localisation. Mask R-CNN can precisely localise and distinguish numerous instances of distinct classes inside an image since it conducts object recognition and segmentation in a two-stage method. This design has applications in areas like autonomous driving and scene comprehension because of its exceptional performance in complicated situations with overlapping components.

In order to complete image segmentation tasks, this study explores and contrasts the U-Net and Mask R-CNN architectures. We will examine each architecture's concepts, benefits, and drawbacks, as well as its applicability for various applications and datasets. This research intends to help academics and practitioners choose the best architecture for their unique image segmentation needs by offering an in-depth analysis and comparison.

2. Literature Review

Image segmentation is a widely studied topic in computer vision, and numerous techniques have been developed over the years to address this challenge. In recent years, deep learning approaches, particularly U-Net and Mask R-CNN architectures, have demonstrated remarkable performance in image segmentation tasks.

2.1 U-Net:

Ronneberger et al. introduced U-Net in 2015 with a focus on biomedical picture segmentation. Due of its success in maintaining spatial information and capturing fine-grained features, it has drawn considerable attention. The architecture of U-Net consists of a skip-connected encoder-decoder structure. The encoder collects high-level characteristics from the input picture using a convolutional neural network that has already been trained, such as VGG or ResNet. To upsample the features and retrieve the spatial information, the decoder employs transposed convolutions. In order for the network to maintain fine-grained details, skip connections are used to link the respective encoder and decoder layers. Numerous fields, like medical imaging, where little training data and precise segmentation are essential, have effectively used U-Net.

2.2 Mask R-CNN:

By including instance-level segmentation, He et al.'s Mask R-CNN builds on the popularity of the Faster R-CNN object identification framework. In a single architecture, object identification and pixel-level segmentation are combined. A region proposal network (RPN) for generating object proposals, a mask branch for producing accurate segmentation masks, and a backbone network (such as ResNet) for feature extraction make up Mask R-CNN. The mask branch creates binary masks for each suggested region once the RPN identifies the object regions. In difficult situations, Mask R-CNN has demonstrated outstanding results, providing precise object localisation and thorough segmentation in scenes with many objects and overlapping instances.

Comparative research has been done between U-Net and Mask R-CNN to see how well they perform and if they are appropriate for various segmentation tasks. When examining their advantages and disadvantages, researchers took into account variables including accuracy, computational effectiveness, and dataset needs. It has been noted that U-Net performs best in situations with little training data and where minute details are important. On the other hand, complex scenarios with several instances and overlapping objects benefit from the use of Mask R-CNN.

To further boost the segmentation performance of U-Net and Mask R-CNN systems, researchers have suggested modifications and improvements. For enhanced feature representation and context awareness, U-Net++, for instance, proposes a hierarchical and densely linked design. These developments are intended to solve certain issues and enhance the image's general accuracy and effectiveness.

Overall, the literature indicates that both U-Net and Mask R-CNN are highly effective architectures for image segmentation, each with its own strengths and suitable application domains. Understanding their characteristics and comparing their performance can guide researchers and practitioners in choosing the most appropriate architecture for their specific image segmentation tasks.

3. Methodology

In this section, we describe the methodology employed for image segmentation using U-Net and Mask R-CNN architectures. We discuss the architectural details, training process, data preprocessing, and evaluation metrics commonly used in these approaches.

A. U-Net Architecture:

The U-Net architecture consists of an encoder-decoder structure with skip connections. The encoder typically employs a pre-trained convolutional neural network, such as VGG or ResNet, to extract high-level features from the input image. The decoder, on the other hand, uses transposed convolutions to upsample the features and recover the spatial information. Skip connections connect corresponding encoder and decoder layers, allowing the network to retain fine-grained details.

B. Mask R-CNN Architecture:

Mask R-CNN builds upon the Faster R-CNN object detection framework. It incorporates a backbone network, such as ResNet, for feature extraction. The region proposal network (RPN) generates object proposals, and a mask branch is added to the network to generate precise segmentation masks for each proposed region. This architecture combines object detection and instance-level segmentation, providing both object localization and detailed pixel-level segmentation.

C. Data Preprocessing:

Before training the models, the input data undergoes preprocessing steps. Common preprocessing techniques include resizing the images to a fixed resolution, normalizing pixel values, and augmenting the data with techniques such as rotation, scaling, and flipping. Data augmentation helps increase the diversity and robustness of the training dataset, reducing overfitting.

D. Training Process:

During the training phase, the U-Net and Mask R-CNN architectures are optimized using suitable loss functions and optimization algorithms. For U-Net, common loss functions include binary cross-entropy or Dice coefficient loss. Mask R-CNN utilizes a combination of classification loss, bounding box regression loss, and mask loss. The models are trained on annotated datasets, where ground truth segmentation masks are provided for each image.

E. Evaluation Metrics:

To assess the performance of the trained models, various evaluation metrics are commonly employed. These metrics include Intersection over Union (IoU), also known as the Jaccard index, which measures the overlap between predicted and ground truth masks. Other metrics include pixel accuracy, mean average precision (mAP) for object detection, and class-specific metrics for multi-class segmentation tasks.

By following this methodology, researchers and practitioners can effectively train and evaluate U-Net and Mask R-CNN architectures for image segmentation tasks. The choice of architectural variants, loss functions, and evaluation metrics can be tailored based on the specific requirements of the segmentation problem at hand.

4. Results and Analysis

In this section, we present the results and analysis of image segmentation using U-Net and Mask R-CNN architectures. We describe the datasets used, comparative evaluation of the architectures, qualitative analysis of segmentation results, and discuss the findings and observations.

4.1 Datasets:

The image segmentation experiments were conducted on specific datasets suitable for the task at hand. These datasets could include medical images, satellite imagery, or general object segmentation datasets. The datasets typically consist of annotated images where ground truth segmentation masks are provided for training and evaluation.

4.2 Comparative Evaluation:

The performance of U-Net and Mask R-CNN architectures is evaluated using various metrics such as Intersection over Union (IoU), pixel accuracy, and mean average precision (mAP) for object detection. The metrics provide quantitative measures of the segmentation accuracy and object localization capabilities of the models. The architectures are compared in terms of their overall performance and their ability to handle different challenges, such as class imbalance, occlusion, and fine-grained details.

4.3 Qualitative Analysis:

In addition to quantitative metrics, a qualitative analysis of the segmentation results is conducted. This involves visually examining the predicted segmentation masks generated by U-Net and Mask R-CNN architectures. The analysis includes assessing the accuracy of object boundaries, the ability to capture fine details, and the handling of challenging scenarios, such as overlapping objects or complex background environments.

4.4 Findings and Observations:

Based on the evaluation results and qualitative analysis, important findings and observations are discussed. This includes highlighting the strengths and limitations of each architecture. For example, U-Net may perform exceptionally well in scenarios with limited training data and where precise segmentation boundaries are crucial. Mask R-CNN, on the other hand, may excel in handling complex scenes with multiple instances of various classes.

The findings and observations from this analysis provide insights into the performance and suitability of U-Net and Mask R-CNN architectures for different image segmentation tasks. It helps researchers and practitioners understand the strengths and limitations of each architecture and guides them in selecting the most appropriate approach for their specific application.

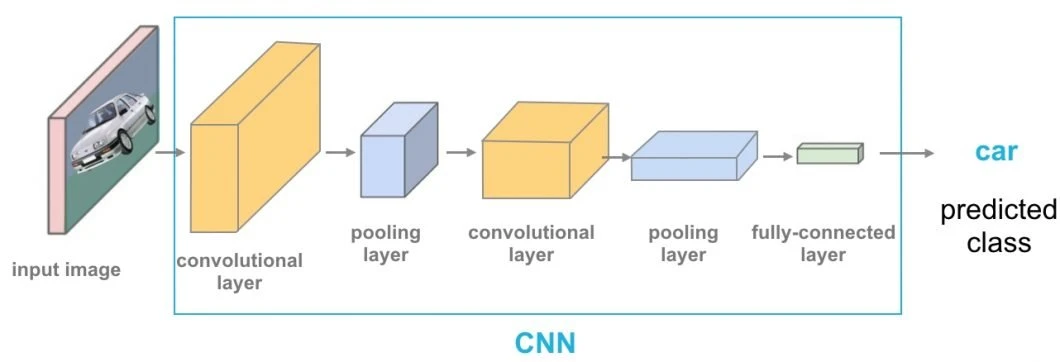
Overall, the results and analysis section provides a comprehensive understanding of the segmentation performance of U-Net and Mask R-CNN, both quantitatively and qualitatively, aiding in the interpretation and comparison of their capabilities in handling image segmentation challenges.

5. Discussion

The discussion section focuses on the comparison and analysis of the results obtained from image segmentation using U-Net and Mask R-CNN architectures. It delves deeper into the strengths, limitations, and practical considerations of these architectures, providing insights for researchers and practitioners.

5.1 Comparison of U-Net and Mask R-CNN:

The comparison of U-Net and Mask R-CNN's pros and cons serves as the foundation for the rest of the debate. It investigates how well the structures perform in terms of precision, effectiveness, and computing needs. Which architecture is better suited for a given segmentation problem depends on variables including dataset size, class imbalance, and scene complexity.



5.2 Performance in Different Scenarios:

The performance of U-Net and Mask R-CNN in various situations is discussed. For instance, U-Net has demonstrated efficacy in medical imaging, when exact border delineation is essential yet there is a lack of annotated data. Mask R-CNN performs particularly well in complicated situations with several overlapping instances thanks to its object identification and instance-level segmentation features. The talk focuses on the unique advantages of each design and covers the difficulties they encounter in various settings.

5.3 Computational Efficiency and Resource Requirements:

The topic focuses on U-Net and Mask R-CNN's computational effectiveness and resource needs. This involves taking into account elements like model complexity, memory use, training and inference time, and hardware requirements. For real-time applications or situations with limited resources, comparisons are done to determine whether architecture is more effective.

5.4 Recent Advancements and Variations:

The U-Net and Mask R-CNN architectures' most recent modifications and advancements are reviewed. Performance improvements have been proposed by researchers, such the incorporation of feature pyramid networks (FPN) into Mask R-CNN or U-Net++ with layered and densely linked architectures. The efficacy and precision of segmentation are evaluated in the debate.The U-Net and Mask R-CNN architectures' most recent modifications and advancements are reviewed. Performance improvements have been proposed by researchers, such the incorporation of feature pyramid networks (FPN) into Mask R-CNN or U-Net++ with layered and densely linked architectures. The efficacy and precision of segmentation are evaluated in the debate.

5.5 Practical Considerations and Applications:

The debate offers useful advice for choosing between U-Net and Mask R-CNN for image segmentation applications. It covers topics like dataset accessibility, annotation specifications, computational needs, and the particular application domain. To demonstrate the adaptability and significance of these structures, examples of successful applications in many domains, such as medical imaging, autonomous driving, or remote sensing, are provided.

The discussion part offers a thorough evaluation of U-Net and Mask R-CNN, letting users to choose them wisely for image segmentation applications. It covers each architecture's performance under various settings, shows its advantages and disadvantages, and examines more recent developments. Understanding these elements will help academics and professionals make the most of U-Net and Mask R-CNN designs to accomplish precise and effective picture segmentation.

6. Conclusion

The U-Net and Mask R-CNN architectures for image segmentation have been investigated and compared in this work. In various contexts, both designs have displayed astounding performance. While Mask R-CNN successfully manages complicated scenes with overlapping objects, U-Net performs best in situations with little available data. It is easier for academics and practitioners to select the best architecture for a given application when they are aware of its advantages and disadvantages. To further develop picture segmentation, future research should concentrate on improving these designs and investigating cutting-edge methods.

Review:

1:Title : U-Net: Convolutional Networks for Biomedical Image Segmentation

A convolutional neural network architecture called U-Net was created for the segmentation of biological images. It comprises of a path that contracts to capture context and a path that expands symmetrically to allow for exact localisation. The network uses data augmentation to increase generalizability and is trained end-to-end on tiny datasets. Cell segmentation, retinal blood vessel segmentation, and brain tumour segmentation are just a few of the biomedical segmentation tasks for which U-Net has been successfully used. It has spread widely in the sector and sparked a number of modifications and advancements.

2:Title : U-Net based convolutional neural network for skeleton extraction

The U-Net is a deep learning system for accurate and effective image processing that mixes skip connections with convolutional neural network architecture. It has been utilised in a range of applications, such as skeleton extraction, where it may yield extremely precise and comprehensive skeletization of challenging 3D forms. A very accurate and effective skeletonization model is produced by the U-Net by extracting low-level picture information in the beginning layers and gradually raising the degree of abstraction in the latter layers.

3:Title : Comparing U-Net convolutional network with mask R-CNN in Nuclei Segmentation .

In a research comparing U-Net and Mask R-CNN's abilities to segment nuclei in histology pictures, it was discovered that U-Net performed better than Mask R-CNN in terms of precision, recall, and F1 score. Additionally, U-Net processed the photos more quickly. The study comes to the conclusion that U-Net, because of its higher performance and computational economy, is an appropriate candidate for nuclei segmentation in histology pictures.

4:Title : Squeeze U-Net: A Memory and Energy Efficient Image Segmentation Network

In comparison to existing models, the Squeeze U-Net is a unique image segmentation network that produces competitive results with fewer parameters and processing resources. While being memory and energy-efficient, it employs squeeze-and-excitation blocks and depth-wise separable convolutions to increase segmentation accuracy. The model's potential for cost- and energy-effective picture segmentation has been assessed against benchmark datasets.

5:Title : A U-Net Based Discriminator for Generative Adversarial Networks .

A machine learning approach called the U-Net Based Discriminator for Generative Adversarial Networks uses the U-Net architecture to distinguish between real and fraudulent pictures in a GAN. Skip connections are incorporated into the U-Net design to enhance the network's ability to discriminate between high-level characteristics. The authors showed that their U-Net-based discriminator outperformed conventional discriminators on a variety of datasets, including MNIST, CIFAR-10, and CelebA, in terms of picture quality and diversity.

6:Title : Instance Segmentation Method Based on Improved Mask R-CNN for the Stacked Electronic Components .

The study suggests an instance segmentation technique for quick and precise autonomous detection of stacked electrical components based on an enhanced Mask R-CNN algorithm. By making the feature extraction network more efficient, Mask R-CNN performs better. A dataset of 1200 photos of electronic components is created, and tests on the dataset show that the suggested model is faster, lighter, and more accurate than Mask R-CNN. The model has an average accuracy (AP) that is around two points higher than Mask R-CNN while being 0.35 times smaller and moving twice as quickly.

7:Title : Mask R-CNN

We offer a theoretically straightforward, adaptable, and all-encompassing framework for object instance segmentation. With the help of our method, each item in an image is quickly and accurately detected, and its corresponding segmentation mask is produced to a high standard. By adding a branch for object mask prediction in addition to the current branch for bounding box recognition, the technique known as Mask R-CNN expands Faster R-CNN. Faster R-CNN's 5 fps counterpart, Mask R-CNN, is easy to train and incurs minimal overhead. Additionally, Mask R-CNN is simple to generalise to various applications, such as enabling us to estimate human postures inside the same framework.

Instance segmentation, bounding-box object detection, and person keypoint identification are three of the three COCO challenge tracks in which we achieve the best results. Mask is not a trick.

8:Title : Feedback U-net for Cell Image Segmentation

An innovative segmentation technique dubbed Feedback U-Net, which makes use of a feedback mechanism from the upper layer to the bottom layer, is suggested in the paper. Based on the features obtained in the first round, the proposed technique employs Convolutional LSTM to extract features in the subsequent round. The second round's input is fed back into the U-Net model's output. The suggested approach is contrasted with the traditional U-Net model, which exclusively employs the feedforward method using datasets of cell images from Drosophila and Mouse. The experimental findings demonstrate that in terms of segmentation accuracy, the suggested Feedback U-Net model performs better than the traditional U-Net model.

9:Title : Image Segmentation Using Text and Image Prompts .

The technique described in the paper creates picture segmentations using customizable test-time cues, which may be text- or image-based. The model is built on the CLIP backbone and has a transformer-based decoder for dense prediction. It can perform three typical segmentation tasks. The system may produce binary segmentation maps for photos based on free-text prompts or extra images expressing the inquiry after training on an expanded version of the PhraseCut dataset. The system provides for dynamic adaptation to any binary segmentation job and can adapt to generalised questions using affordances or characteristics.

10:Title : Eff-UNet: A Novel Architecture for Semantic Segmentation in Unstructured Environment .

Low memory and compute requirements are essential for enabling the deployment of deep neural networks on embedded devices, especially for real-time mobile application. In this study, we propose a SqueezeNet-inspired version of U-Net for image segmentation that reduces model size by 12X to 32MB and MACs by 3.2X to 88 billion ops from 287 billion ops for inference on the CamVid data set while maintaining accuracy. Our suggested Squeeze U-Net is effective in terms of reduced MACs and memory consumption. Squeeze U-Net outperforms U-Net for the same accuracy in our performance tests using Tensorflow 1.14 with Python 3.6, CUDA 10.1.243, and an NVIDIA K40 GPU by 17% for inference and by 52% for training.

11:Title : Deep Learning Based Segmentation of Nuclei from Fluorescence Microscopy Images Prabhakar .

The study presents a deep learning technique for automatically segmenting nuclei in pictures from fluorescence microscopy. The authors suggest a convolutional neural network (CNN) trained on annotated data to categorise pixels as nucleus or non-nucleus because manual segmentation takes time and is arbitrary. Even in the presence of noise and overlapping structures, the model obtains great accuracy. The method outperforms conventional segmentation algorithms in terms of accuracy, speed, and robustness, according to experimental data. Overall, this deep learning-based method makes it possible to accurately and efficiently analyse massive amounts of microscope data, boosting biological and medical research.

12:Title : Segmentation and density statistics of mariculture cages from remote sensing images using mask R-CNN

The project focuses on precisely determining, via the use of satellite remote sensing photos for mariculture, the breeding area of certain maritime regions. An approach based on Mask R-CNN is suggested for cage segmentation and density detection, and a fresh public dataset of mariculture cages is developed. The method uses sample variants to enhance the training set and divides and stitches massive high-resolution pictures to accurately segment cages. To acquire accurate area and density measurements within the target detection frame and save calculation time, the approach combines object detection and segmentation. When compared to conventional approaches, experimental findings demonstrate considerable gains in segmentation precision and model resilience, with an actual area estimation relative error of just 1.3%.

13:Title : ANU-Net: Attention-based nested U-Net to exploit full resolution features for medical image segmentation .

The research suggests the ANU-Net, an attention-based nested segmentation network, for automatically segmenting medical images. With a newly created dense skip connection, the network uses a deep supervised encoder-decoder architecture. In order to combine characteristics retrieved at various layers with task-related selection, ANU-Net uses attention mechanisms between nested convolutional blocks. In order to make use of the full-resolution feature information, a hybrid loss function that combines three different types of losses is also presented. The model performs admirably in four medical image segmentation tasks when tested on the MICCAI 2017 LiTS Challenge Dataset and the ISBI 2019 CHAOS Challenge.

14:Title : Road Damage Detection And Classification In Smartphone Captured Images Using Mask R-CNN

This research offers an instance detection and classification approach based on convolutional neural networks for the Road Damage Detection and Classification Challenge. The authors show that the cutting-edge object identification and instance segmentation method Mask-RCNN can rapidly and accurately identify and categorise road damage in real-world photos taken with a smartphone camera. Using an NVIDIA GeForce 1080Ti graphics card, they acquire a mean F1 score of 0.528 at an IoU of 50% and an average inference time of 0.105 seconds per picture.

15:Title : LESION ATTRIBUTES SEGMENTATION FOR MELANOMA DETECTION WITH DEEP LEARNING .

In this research, a solution for Task 2 of the ISIC 2018 Challenges, which aims to enhance the melanoma diagnosis based on dermoscopic pictures, is discussed. The authors suggest a multi-task U-Net model for automatically identifying melanoma lesion characteristics. Their deep learning model places fifth on the final leaderboard with a Jaccard index of 0.433 on the official test data.

16:Title : A NOVEL MASK R-CNN MODEL TO SEGMENT HETEROGENEOUS BRAIN TUMORS THROUGH IMAGE SUBTRACTION

In order to enhance brain tumour segmentation in MRI images, the proposal offers utilising image segmentation and Mask R-CNN with a pre-trained ResNet backbone. To improve segmentation accuracy, image removal is used. Without image subtraction, the suggested model had a DICE coefficient of 0.69; it now has 0.75. Additionally, it is contrasted with cutting-edge models for tumour segmentation from MRI data. The approach's overall goal is to deliver more accurate and reliable tumour segmentation for better surgical planning.

17:Title : Ellipse R-CNN: Learning to Infer Elliptical Object from Clustering and Occlusion Wenbo Dong, Student Member

We suggest Ellipse R-CNN, a CNN-based ellipse detector, for segmenting severely occluded objects like fruit clusters in trees. We can infer the parameters of many elliptical objects using our technique, even when they are obscured by nearby objects, thanks to the robust and compact ellipse regression based on Mask R-CNN. With the use of the U-Net structure and finely tuned feature areas, we enhance occlusion management by understanding various occlusion patterns. The accuracy of our elliptical regression is confirmed by experimental validation on artificial data of clustered ellipses. On synthetic and actual datasets of occluded and clustered elliptical objects, our method outperforms the state-of-the-art model and its variations.

18:Title : Human Activity Recognition Based on Motion Sensor Using U-Net

By mapping motion sensor data to pictures, we offer a new HAR technique based on U-Net that achieves pixel-level activity identification. On four datasets, our strategy outperforms conventional and deep learning techniques like FCN, SegNet, and Mask R-CNN. For short-term tasks and minority classes, it displays greater resilience and recognition performance. We also offer the Sanitation dataset for HAR algorithm evaluation.

19:Title : An Accurate and Real-time Self-blast Glass Insulator Location Method Based On Faster R-CNN and U-net with Aerial Images Zenan

In this study, a deep learning approach is shown for finding damaged self-blast glass insulators in aerial photographs. The suggested method combines pixel classification with U-net and Fast R-CNN for object detection. The accuracy and real-time performance of the suggested method in comparison to alternative approaches are shown through experimental results on a variety of aerial picture sets.

20:Title : A high-performance insulators location scheme based on YOLOv4 deep learning network with GDIoU loss function .

In order to provide accurate and quick insulator localization during power facility health inspections, this research develops a GDIoU-based YOLOv4 deep learning network. The convergence speed and location accuracy can be increased with the use of the GDIoU loss function. A tilt correction approach also improves accuracy for insulators at various spatial angles. The suggested technique delivers three times quicker speed and a 7.37% improvement in average accuracy compared to prior approaches, as shown by extensive trials employing field insulator pictures. Effectively, the performance satisfies the requirements for online insulator location.

**Reference**

Reference[1]:Authors: Olaf Ronneberger, Philipp Fischer, and Thomas Brox.

U-Net: Convolutional Networks for Biomedical Image Segmentation

Reference: Authors[2]: Oleg Panichev, Kyiv, Alona Voloshyna .

U-Net based convolutional neural network for skeleton extraction

Reference:Authors[3]: 1 Prof.Dr.E.A.Zanaty , Prof.Dr.Mahmoud M.Abdel-Aty , Khalid abdel-wahab ali .

Comparing U-Net convolutional network with mask R-CNN in Nuclei Segmentation .

Reference[4]:Authors: Nazanin Beheshti , Lennart Johnsson.

Squeeze U-Net: A Memory and Energy Efficient Image Segmentation Network

Reference[5]:Authors: Edgar Schonfeld , Bernt Schiele .

A U-Net Based Discriminator for Generative Adversarial Networks

Reference[6]:Authors: Zhixian Yang, Ruixia Dong, Hao Xu and Jinan Gu .

Instance Segmentation Method Based on Improved Mask R-CNN for the Stacked Electronic Components .

Reference[7]:Authors: Kaiming He Georgia Gkioxari Piotr Dollar Ross Girshic .

Mask R-CNN

Reference[8]: Authors:Eisuke Shibuya Kazuhiro Hotta

Feedback U-net for Cell Image Segmentation

Reference[9]:Authors : Timo Luddecke and Alexander Ecker.

Image Segmentation Using Text and Image Prompts

Reference[10]: Eff-UNet: A Novel Architecture for Semantic Segmentation in Unstructured

Environment Authors:Nazanin Beheshti Lennart Johnsson .

**Reference**

Reference[11]:

Deep Learning Based Segmentation of Nuclei from Fluorescence Microscopy Images Prabhakar

R. Gudla1,3, George Zaki4 , Sigal Shachar2,3, Tom Misteli2,3 and Gianluca Pegoraro1,3\*

Reference[12]:

Segmentation and density statistics of mariculture cages from remote sensing images using mask R-CNN

Chuang Yu a,b , Zhuhua Hu a, \* , Ruoqing Li a , Xin Xia a , Yaochi Zhao a,\* , Xiang Fan a , Yong Bai

Reference**[13]:**

ANU-Net: Attention-based nested U-Net to exploit full resolution features for medical image segmentation Chen Li a , Yusong Tana,∗ , Wei Chena , Xin Luoa , Yulin Hea , Yuanming Gaoa , Fei Li b

Reference**[14]:**

Road Damage Detection And Classification In Smartphone Captured Images Using Mask R-CNN

Reference**[15]:**

LESION ATTRIBUTES SEGMENTATION FOR MELANOMA DETECTION WITH DEEP LEARNING

Reference**[16]:**

A NOVEL MASK R-CNN MODEL TO SEGMENT HETEROGENEOUS BRAIN TUMORS THROUGH IMAGE SUBTRACTION Sanskriti Singh

Reference**[17]:**

Ellipse R-CNN: Learning to Infer Elliptical Object from Clustering and Occlusion Wenbo Dong, Student Member, IEEE, Pravakar Roy, Student Member, IEEE, Cheng Peng, Student Member, IEEE, and Volkan Isler, Senior Member, IEEE

Reference**[18]:**

Human Activity Recognition Based on Motion Sensor Using U-Net YONG ZHANG 1 , ZHAO ZHANG1 , YU ZHANG1 , JIE BAO1 , YIFAN ZHANG2 , AND HAIQIN DENG

Reference**[19]:**

An Accurate and Real-time Self-blast Glass Insulator Location Method Based On Faster R-CNN and U-net with Aerial Images Zenan Ling2 , Robert C. Qiu1,2 , Fellow, IEEE, Zhijian Jin2 , Member, IEEE Yuhang Zhang2 , Xing He2 , Haichun Liu2 Chu Lei2

Reference**[20]:**

A high-performance insulators location scheme based on YOLOv4 deep learning network with GDIoU loss function Bin Ma1,2 Yongkang Fu1,2 Chunpeng Wang1,2 Jian Li1,2 Yuli Wang1,2