Journal of Real-Time Image Processing (2025) 22:39
https://doi.org/10.1007/s11554-024-01615-5

RESEARCH

PL-UNet: a real-time power line segmentation model for aerial images based on adaptive fusion and cross-stage multi-scale analysis

Qian Zhao[1] · Haosheng Fang[1] · Yuye Pang[1] · Gehan Zhu[1] · Zhengzhe Qian[1]

Received: 26 November 2024 / Accepted: 21 December 2024 / Published online: 7 January 2025

Abstract
The examination of power transmission lines using UAVs (Unmanned Aerial Vehicles) is crucial for ensuring grid security. However, it is challenging for existing deep learning models to achieve a balance between accuracy and efficiency in recognizing power lines, especially when they are affected by intricate environmental backdrops and the thin structure of power lines. To address this issue, this paper proposes an improved model based on U-Net (PL-UNet), which aims to improve the ability of UAVs to recognize power lines in complex environments in real time. To reduce the model parameters, the lightweight EfficientNetV2-S is chosen as the encoder. To address the issues of information redundancy and insufficient local structure capture caused by the skip connections, we propose a multi-scale attention gate (MSAG) in the decoding part to improve the accuracy of key region feature extraction with less computational cost. Meanwhile, the dynamic weighted fusion (DWF) module is designed to effectively fuse the features through adaptive weighting to improve the flexibility of feature expression. After completing feature fusion, we further introduce a lightweight cross-stage partial pyramid block (CPPB) module, which performs multi-scale enhancement and channel optimization of the fused features through integrating multi-scale convolutional operations and separating and fusing feature channels. Finally, the hybrid loss function of weighted cross-entropy and dice is used to solve the category imbalance problem. Comparison experiments and ablation analysis are performed on top of the power line public dataset. The proposed PL-UNet has achieved 79.98% mIoU, with a parameter count of 21.27M and a detection speed of 56.86 fps. This shows that the network has a good real-time segmentation performance.

Keywords Power line segmentation · U-Net · Adaptive feature fusion · EfficientNetV2

1 Introduction

Power lines are the only way to deliver electricity, so regular inspection of transmission lines is crucial to guarantee the secure functioning of the electrical network. Transmission lines are often laid in environments with complex and variable terrain. This renders the conventional manual inspection approach not merely tedious and resource-demanding, but also deficient in efficiency, coupled with inherent safety hazards. With the development of artificial intelligence and smart grid, drones [1, 2] are gradually replacing the traditional manual inspection as a new way of power line inspection. However, power lines are a major obstacle in overhead power transmission systems. The very thin structure of power lines may lead to accidents caused by collisions that are not recognized by UAVs. Therefore UAV inspections must precisely avoid these lines with thin structures to ensure safe flight. Accurate transmission line segmentation technology is important for UAVs to recognize power lines, complete inspection tasks, and advance the progress of navigational technology and intelligent inspection systems.

Presently, the methods for segmenting power lines are predominantly categorized into two classes: the classical image processing techniques and the approaches grounded in deep learning. For classical image processing techniques, Azevedo et al. [3] proposed a LiDAR-based method (PL2DM) for detecting power lines in challenging environments by adaptively analyzing point clouds and leveraging power line features. Senthilnath et al. [4] proposed a bilevel segmentation framework utilizing metacognitive strategies for power line detection in UAV images, enhancing accuracy with color and shape cues. Zhao et al. [5] employed adaptive Gaussian pyramid and multi-scale LSD for power

✉ Yuye Pang
846715668@qq.com

[1] School of Electronics and Information Engineering, Shanghai University of Electric Power, Shanghai, China

line location identification followed by feature extraction with OMRF and WRAG techniques, and tracking with Kalman filtering and least squares. Wang et al. [6] proposed a complex environment segmentation method, starting with ground simulation via RANSAC, converting 3D point clouds to 2D for line detection, and employing Euclidean clustering for power line detection. Zhao et al. [7] improved power line detection in complex settings by using morphological filtering and multi-scale analysis to identify high-contrast linear structures.

Above all, traditional image processing methods boast high computational efficiency in the extraction of power lines. However, these methods lack portability and robustness across different scenes. Particularly when dealing with real-time tasks, such as aerial imagery, their performance often falls short of the required standards. In contrast to traditional image processing techniques, deep learning-based approaches have demonstrated superior generalization capabilities and are now extensively applied across a variety of computer vision tasks. Consequently, numerous deep learning-powered end-to-end semantic segmentation frameworks have been introduced to effectively segment power lines from aerial imagery. Among them, Gao et al. [8] introduced a high-performance parallel branching architecture for instantaneous segmentation of aerial power lines. This network uses an AFDB module to extract short-range features. The network fuses long-range features through skip connections. However, it has limitations with complex backgrounds due to low-resolution training data. Yang et al. [9] created an encoder-decoder network for power line detection in aerial images. They added attention and multi-scale feature fusion modules to address class imbalance and context processing. Adding these two modules improves the model's performance but also increases computation, slowing down detection, making it less suitable for applications that need quick responses. Choi et al. [10] proposed a multi-modal image feature fusion module that combines visible and infrared images to improve the segmentation performance of power lines. However, this fusion and attention mechanism may increase the computational complexity, affect the real-time performance, and be sensitive to the shooting angle. Yang et al. [11] proposed PLE-Net, an innovative network for power line detection, employing self-attention and multi-scale features to address class imbalance. Despite challenges from complex backgrounds, there is scope for enhancing the model's segmentation precision. Abdelfattah et al. [12] proposed a Power Line Segmentation Model (PLGAN) based on Generative Adversarial Networks, specifically for aerial images. The model is able to capture scene context, shape features and visual features through adversarial training to achieve more accurate semantic segmentation. However, this method has some limitations and its application is difficult to be extended to other domains. Chen et al. [13] proposed PL-Deeplab, an efficient multi-branch cascade network. They designed an MCNet designed to enhance the expressive power of the model, and introduced an OSAFP to enhance the global consistency between pixels. Nevertheless, the model's extensive parameter set can limit its applicability in environments with limited resources or in scenarios requiring real-time performance.

Although deep learning-based methods outperform traditional image processing methods in terms of performance, extracting power lines from aerial images is still a challenge due to a variety of factors. First, the background of aerial images is complex and varied, and covers a wide range of areas. The extraction of power lines is easily affected by factors such as roads, vegetation, buildings, railings, weather, etc. The similarity in shape and color between these background elements and power lines may lead to misclassification or inaccurate segmentation. Secondly, power lines are relatively narrow and account for a minimal portion of the image's pixels. So, this makes them difficult to be segmented efficiently from complex backgrounds. To address these issues, we used the TTPLA public dataset provided by Abdelfattah et al. [14] for our study. This dataset not only contains features such as unbalanced data categories [15] and structural complexity, but is also highly challenging. Utilizing this dataset allows us to more effectively assess and refine the model, which aims to improve the accuracy and reliability of extracting power lines from intricate environments. Combining the above analysis, we propose a U-Net-based improved power line segmentation model. Our main contributions include the following:

1. The introduced multi-scale attention gate (MSAG) module enhances the model's capacity for detecting local patterns within power line scenes. This is achieved by expanding the sensory field through the use of lightweight convolutions and concentrating on features from critical regions.
2. The dynamic weighted fusion (DWF) module in the decoder flexibly fuses the features from the skip connections and the up-sampled features through an adaptive weighting mechanism. It enhances the robustness of the feature expression and realizes the accurate segmentation of the power line boundary.
3. The designed cross-stage partial pyramid block (CPPB) significantly boosts the model's capacity to detect fine details by capturing features across multiple scales and utilizing an efficient cross-stage partial (CSP) framework, while improving the segmentation accuracy in complex background scenes.
4. Visual image experiments demonstrate that the approach outlined in this document excels in both the precision

and speed of power line extraction from aerial imagery, while maintaining an optimal balance between segmentation accuracy and model complexity.

2 Related work

2.1 U-Net

U-Net [16] network proposed by Ronneberger et al. is the classic encoder-decoder segmentation framework and one of the dominant structures, which has shown excellent performance on different segmentation scenes. It mainly consists of an encoder and a decoder. For the encoder, each unit consists of two 3×3 convolutions and a maximum pooling layer for feature downsampling, where the maxpooling layer has a step size of 2. After each convolutional layer, a batch normalization step and a ReLU activation function are applied sequentially. For the decoder, it consists of a transposed convolution or bilinear interpolation for feature upsampling and two 3×3 convolutions. Concurrently, skip connections are established between the encoder and decoder, facilitating the direct transfer of the encoder's feature map to the corresponding decoder layer. Within each upsampling phase of the network, these decoder feature maps are merged with those from the encoder at corresponding spatial locations along the channel dimension. Such a design efficiently retains spatial details and minimizes feature loss from repeated convolutions and pooling, substantially enhancing segmentation precision. Nonetheless, employing it for detecting power lines using UAVs reveals several challenges, including inadequate feature extraction, a restricted perceptual scope, and poor integration of multi-level information. Consequently, leveraging the U-Net model as a foundation, we have developed a network specifically for inspecting power lines called PL-UNet.

2.2 Efficient backbone network

Convolutional neural networks are good at capturing spatial relationships in images, and thus are often used as a key part of an encoder. VGG [17] lays a stable foundation for the development of CNNs by increasing the depth of convolutional layers and extracting richer features layer-by-layer. ResNet [18] introduces residual connectivity mechanism, which successfully solves the problem of gradient vanishing in the training of deep networks. Xception [19] improved the Inception [20] architecture by replacing the original module with deep separable convolution, which further improved the computational efficiency and feature extraction capability. For resource-constrained mobile devices, MobileNets [21] utilizes depth-separable convolution to enable CNNs to operate efficiently with limited computational resources. EfficientNet [22] proposes a composite scaling strategy, which builds a more efficient model architecture by balancing the depth, width, and resolution of the network. GhostNet [23] introduces the Ghost module, which optimizes the performance of lightweight networks by generating more feature mappings to reduce the number of parameters and computation. In order to improve the feature extraction capability and model efficiency of the encoder of U-Net, we adopt EfficientNetV2-S as the encoder of U-Net. Tan et al. [24] proposed EfficientNetV2 with faster training speed and better parameter efficiency, and its structure combines the Fused-MBConv and the MBConv modules in EfficientNetV1. Compared with other mainstream network structures, such as Vision Transformers, ResNet, DenseNet and other networks, EfficientNetV2 shows significant advantages in speed and parameter efficiency, effectively shortening the training time and reducing the consumption of computational resources. Meanwhile, EfficientNetV2 shows excellent performance in image classification tasks. In particular, its balance between speed and accuracy is significantly improved when processing complex datasets. This model can be flexibly integrated into the U-Net architecture, which effectively addresses the challenge of vanishing gradients in deep networks by adopting techniques such as progressive learning strategy and extended residual connectivity. Acting as the encoder, EfficientNetV2 substantially boosts the model's feature extraction performance.

2.3 Efficient convolution module

In the realm of semantic segmentation, the decoder must incrementally enhance the resolution of the feature map and precisely reconstruct the boundaries and fine details of the target objects. So, the design of an efficient convolution module in the decoder is crucial. U-Net uses a convolutional module (DoubleConv) in the decoder section, which consists of two 3×3 standard convolutional operations to enhance the fusion feature of up-sampled features and features from the skip connections. In this process, Batch Normalization (BN) and ReLU activation functions are added to improve the generalization. However, this design has some limitations in terms of computational efficiency and feature fusion capability. CSPNet [25] introduces a channel segmentation and fusion mechanism, which enriches the feature representation and reduces the computation by dividing the input features into two paths to be processed separately and then fused. ELAN [26] employs a multi-path feature aggregation design, which extracts detailed and global information at different scales, enhancing the module's ability to capture features of complex scenes. CaraNet [27] introduces a convolution with different expansion rates to extract multi-scale features, and uses layer-by-layer accumulation for feature fusion. It

effectively improves the segmentation capability of small targets and complex scenes. LFPNet [28] introduces a linear feature perception convolution group (LFPCG), which introduces the ability of multi-scale feature capture through a multi-branch convolutional design, including standard convolution, expansion convolution, and asymmetric convolution. While taking into account both the parameter efficiency and the computational complexity, it can efficiently capture the contextual information at different scales. Meanwhile, Liu et al. [29] adopted a semantic flow alignment module (SFAM), which focuses on solving the semantic misalignment problem in the process of high- and low-resolution feature fusion. SFAM aligns high and low-resolution features by introducing semantic offset domains, thus preserving the details in recovering the high-resolution feature maps. It significantly improves the ability of boundary reconstruction and detail restoration, which is especially suitable for the scenes with weak contrasts and complex backgrounds. To tackle the challenges posed by intricate background elements within aerial images, we have proposed an efficient Cross-Stage Partial Pyramid Block (CPPB) that is capable of effectively extracting and merging features.

3 Segmentation framework

As shown in Fig. 1, the PL-UNet proposed in this paper is based on an encoder–decoder architecture, where the encoder part is used to extract features and the decoder part is used to recover the spatial resolution. Specifically, EfficientNetV2-S is introduced in the encoder for efficient feature extraction, which is divided into five stages, from stage1 to stage5. Three modules are proposed in the decoder part: Multi-scale Attention Gate, Dynamic Weighted Fusion, and Cross-stage Partial Pyramid Block. Among them, Multi-scale Attention Gate is used to optimize the feature representation. Dynamic Weighted Fusion module is responsible for fusion between the skip connections and the up-sampling layer. Cross-stage Partial Pyramid Block efficiently extracts the multi-scale features. The design of decoder module is described in detail below.

3.1 Multi-scale attention gate

Skip connections are prone to information redundancy and noise in U-Net. This can cause the decoder to process too many irrelevant features, making the model overly reliant

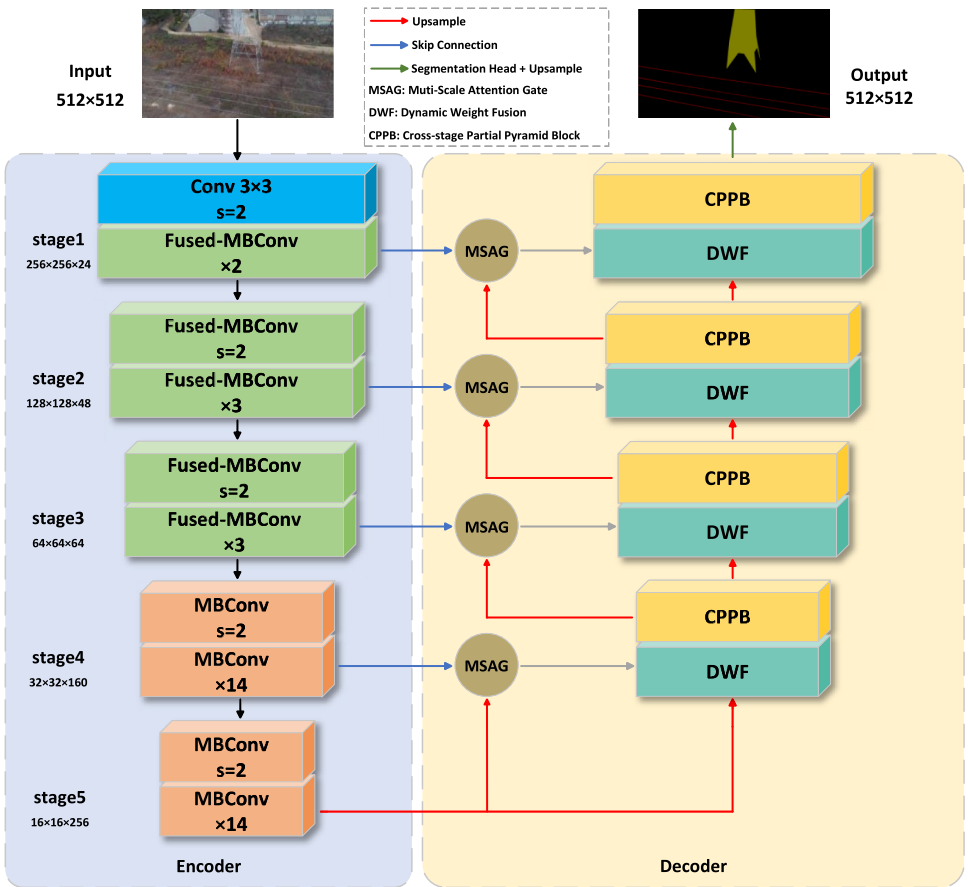

Fig. 1 Entire structure of PL-UNet

on details at the expense of the overall structure. Attention U-Net [30] improves the model's focus on critical regions by introducing an attention mechanism to filter important features and reduce the transfer of irrelevant information. It uses 1×1 convolution to process the gating signal g(up-sampled features) and input feature mapping x(features from skip connections) separately, but its limitation is that it lacks spatial receptive fields to capture local structural features efficiently. SUnet [31] employs Efficient Attention Gate (EAG), which utilizes group convolutions and residual connections to achieve efficient capture of important features and reduce computational complexity. EMCAD [32] introduces Large-kernel Grouped Attention Gate(LKAG) to capture a larger local context through large-kernel group convolution to capture important features with low computational cost. Inspired by both, we propose a new Multi-Scale Attention Gate(MSAG) to effectively capture local structural features through a wider sensory field to enhance the accuracy of the model in power line scene segmentation.

Refer to Fig. 2 for the structure, the 3×3 Group Convolution is used to replace the 1×1 convolution in the Attention U-Net model, thereby increasing the receptive field and enhancing the ability to extract spatial features. The two 3×3 grouped convolutions are denoted by $GConv_g$ and $GConv_x$ for g and x, respectively. After grouped convolution, these convolved features are normalized by batch normalization(BN) and merged by element-by-element addition. The generated feature map is activated by activation function($ReLU$) to get $\psi_1(g,x)$. Next, the above feature maps activated by $ReLU$ are inputted into three convolutional layers with different scales: 1×1 standard convolution, 3×3 depth-wise convolution, and 5×5 depth-wise convolution, respectively. The 1×1 standard convolution is mainly used to aggregate global information. And the 3×3 depth-wise convolution and the 5×5 depth-wise convolution are used to extract multi-scale local features through convolutional kernels with different receptive fields, respectively. The feature maps produced by the three convolutions are fused by element-by-element addition to combine the information at each scale and aggregate features at different levels to enhance the feature representation capability of the model. Finally, the fused feature maps are subjected to feature compression and normalization by 1×1 standard convolution and BN. Then a single-channel attention weight map is generated by a Sigmoid activation function to obtain $\psi_2(g,x)$. The weight map multiplies with the input features element-wise to yield the final output $F_{MSAG}(g,x)$, boosting the model's focus on key areas for more efficient salient region detection. The formula is shown in Eqs. 1–3.

$$\psi_1(g,x) = ReLU(BN(GConv_g(g) + BN(GConv_x(x)))) \tag{1}$$

$$\psi_2(g,x) = \sigma(BN(C_1(\psi_1(g,x)) + DWC_3(\psi_1(g,x)) + DWC_5(\psi_1(g,x)))) \tag{2}$$

$$F_{MSAG}(g,x) = x \cdot \psi_2(g,x) \tag{3}$$

where σ represents to Sigmoid activation function, C_n represents to n×n standard convolution, and DWC_n represents to n×n depth-wise convolution.

Due to the lightweight design of depth-wise convolution and grouped convolution, the computational overhead is reduced while a larger spatial context is captured. And the multi-scale feature fusion enhances the expressive ability of the model, which can better activate relevant features and suppress irrelevant features.

3.2 Dynamic weighted fusion

In the standard U-Net decoder, shallow features in the encoder are spliced with deep features in the decoder by skip connections to combine deep semantic information with shallow detail information. However, this simple splicing approach may lead to the problems of information redundancy and fusion imbalance. This is because feature maps at different layers may contain duplicate information, resulting in poor segmentation of the boundaries of power lines. To solve the above problems, the Dynamic Weighted Fusion (DWF) is designed in this paper. This module enhances the model's performance in fine-grained target segmentation by gradually fusing feature maps of different resolutions through adaptive weighting. It makes the feature representation more flexible and robust and effectively integrates deep semantic information and shallow detail information.

The up-sampled feature F_d and the MSAG-processed feature F_{MSAG} are used as inputs to the DWF, where F_{MSAG} is feature-mapped by a single 3×3 standard convolution operation to obtain F_l. As shown in Fig. 3, which significantly enhances the edge information of the image. The DWF splices F_l and F_d in channel dimensions. With 1×1 convolution and Sigmoid function, we can generate the weight map F_{Concat} from the spliced feature map. This weight map

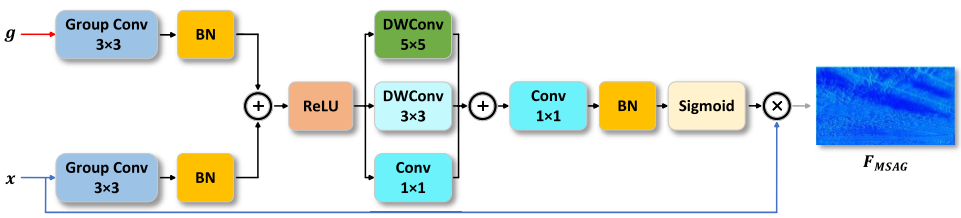

Fig. 2 Basic structure of multi-scale attention gate

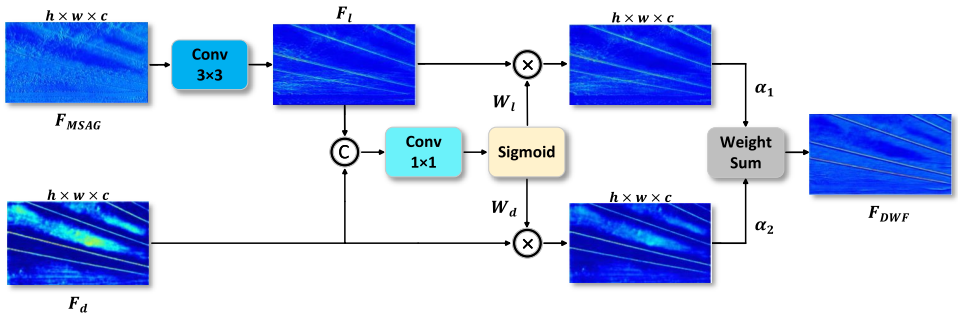

Fig. 3 Basic structure of dynamic weighted fusion

is used to adaptively weight the features with different resolutions, thus controlling the proportion of their contribution to the fusion process. The formula is shown in Eqs. 4–5.

$$F_l = Conv_{3\times3}(F_{MSAG}) \tag{4}$$

$$F_{Concat} = \sigma \cdot Concat(F_l, F_d) \tag{5}$$

where σ indicates the Sigmoid function and *Concat* indicates the feature splicing with 1×1 convolutional extraction operation.

The spliced fused feature F_{Concat} is decomposed into W_l and W_d along the channel dimension, which correspond to the adaptive weights of F_l and F_d, respectively. The weighted representations are used to generate shallow and deep features. It can be observed that the weighted features are able to focus on the edge region of the power line more effectively. It indicates that the module fully utilizes the global background information of F_l and the local detail information of F_d to achieve more accurate feature fusion. Finally, two learnable parameters α_1 and α_2 are used to linearly combine the weighted F_l and F_d to generate the final fused feature F_{DWF}. The formula is shown in Eq. 6.

$$F_{DWF} = \alpha_1 \cdot (F_l \cdot W_l) + \alpha_2 \cdot (F_d \cdot W_d) \tag{6}$$

This approach dynamically controls the feature contributions of different resolutions through adaptive weights and learnable parameters. It can effectively balance the global semantic information and local detail information, which improves the performance of feature fusion.

3.3 Cross-stage partial pyramid block

In U-Net decoder, the DoubleConv module, which consists of two consecutive 3×3 convolutional layers, is responsible for feature extraction and nonlinear mapping. It is usually used in conjunction with the batch-normalized BN and ReLU activation functions. However, the DoubleConv module also has certain limitations, such as high computational overhead, missing feature selection mechanism, insufficient gradient mobility, etc. This limits its performance in more complex segmentation tasks such as power line scenes. Therefore, we propose the Cross-stage Partial Pyramid Block (CPPB), which improves the feature expression capability, computational efficiency and gradient mobility with the help of feature channel segmentation and multi-path convolution. So, it is more suitable for multi-scale feature extraction and complex segmentation tasks.

As shown in Fig. 4, the input feature F_{DWF} undergoes a 1×1 convolution to uniformly divide the cross-channel features into two parts, X_1 and X_2. Among them, X_1 undergoes a 1×1 convolution to retain the global information; and X_2 extracts the local detail information through two cascaded 3×3 convolutional layers. Subsequently, the local feature representation of X_2 is enhanced by effectively fusing the local and global information through element-by-element summation of the detailed features of X_2 with the global information of X_1. The dual-path design guarantees the retention of global information and enhances the capture of local spatial details. Finally, the features of the two paths are merged by a channel splicing operation to fuse the global

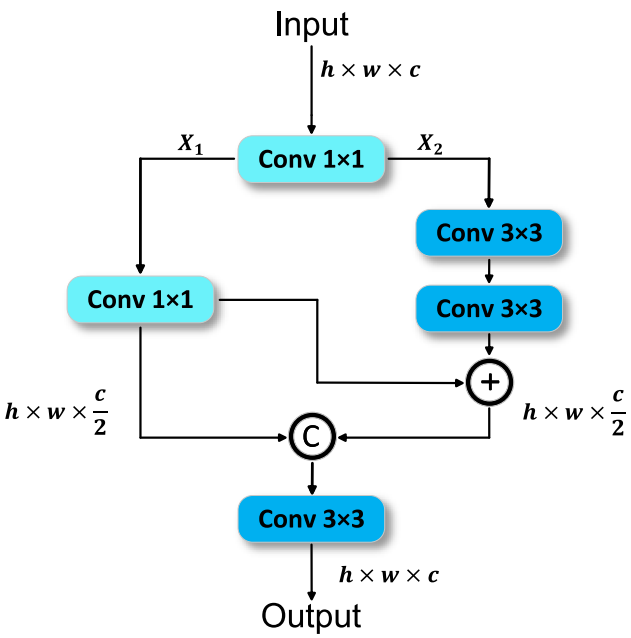

Fig. 4 Basic structure of cross-stage partial pyramid block

and detailed features. The final features F_{CPPB} are then output through a 3×3 convolutional layer combining BN and SiLU [33] activation functions. The path information is further integrated and the nonlinear representation is enhanced. The computational formula is shown in Eq. 7.

$$F_{CPPB}(F_{DWF}) = C_3(Concat(C_1(X_1), C_3(C_3(X_2)) + C_1(X_1))) \tag{7}$$

where C_n refers to the n×n standard convolution.

3.4 Loss function

Loss functions are instrumental in the training of neural networks to achieve optimal convergence. For image segmentation tasks that require pixel-wise classification, the Cross-Entropy (CE) loss function is often the preferred choice. Since the TTPLA dataset contains multiple categories and the number of categories varies widely. So, weighted cross entropy loss [34] is used to solve the problem of category imbalance as shown in Eq. 8.

$$Loss_{CE} = -\sum_{i=1}^{N} w_i \cdot t_i \cdot \log\left(p_i\right) \tag{8}$$

where t_i indicates the corresponding ground truth, p_i indicates the predicted value, and w_i indicates the weight.

In the context of aerial image analysis, addressing the challenge of segmenting power lines is critical due to their slender nature and the prevalence of complex backgrounds. The weighted cross-entropy loss function often underperforms in scenarios with class imbalance, as power lines constitute a minor portion of the overall image pixels in aerial photography. In comparison, the dice loss [35] evaluates the performance of the network by measuring the similarity between the predicted results and the real labels. This approach helps to solve the problem of category imbalance and overfitting. Because it focuses on the overlap between predicted and real data rather than simply calculating classification errors. Dice loss ($Loss_{Dice}$) is obtained by the dice coefficient (*DiceCoefficient*). The formula for the dice coefficient is shown in Eq. 9.

$$DiceCoefficient = \frac{2|X \cap Y|}{|X| + |Y|} \tag{9}$$

where X is the predicted value, Y is the ground truth, $|X \cap Y|$ indicates the intersection of predicted value and the ground truth. This leads to the formula for calculating Dice loss, which is shown in Eq. 10.

$$Loss_{Dice} = 1 - DiceCoefficient \tag{10}$$

Finally, the weighted cross-entropy loss and the dice loss are combined to train the model more efficiently, as shown in Eq. 11.

$$Loss = Loss_{CE} + Loss_{Dice} \tag{11}$$

4 Experiments

4.1 TTPLA dataset

To verify the effectiveness of PL-UNet in power line segmentation, we evaluated the network using the TTPLA dataset. The TTPLA dataset comprises approximately 1200 images, each featuring a high resolution of 3840×2160. Each image includes detailed annotations for power lines and transmission towers. In this study, the dataset was segmented in an 8:1:1 ratio, allocating 958 images to training, 119 to validation, and 126 to testing.

4.2 Experimental setting

The PL-UNet network is constructed with the Pytorch 2.1.0 framework and validated on a GeForce RTX 4080 Laptop GPU platform with 12GB of memory. All the experiments comparing the networks were conducted in the same hardware environment and with the same dataset. For network training, the SGD optimizer was applied to the TTPLA dataset, specifying a batch size of 4, an initial learning rate of 7e-3, a weight decay of 1e-4. The total training epoch is set as 100, and a polynomial decay strategy (poly) with preheating was used to adjust the learning rate. The encoder part was initialized using the pretrained weights of EfficientNetV2-S on ImageNet-1K and further fine-tuned. Finally, the images were resized to 512 × 512, and several data enhancement methods were used. The images were first scaled randomly with scaling ratios ranging from 25% to 200%. The images were flipped horizontally with a 50% probability. In addition, the image has a 25% chance of being rotated with a rotation angle between -5° and 5°. Color adjustments include random variations of hue, saturation, and brightness with adjustments of 0.2, 0.8, and 0.4, respectively.

4.3 Evaluation indicator

For a comprehensive evaluation of the PL-UNet's performance, this research utilizes several key metrics: mean Intersection over Union (mIoU), mean Pixel Accuracy (mPA), Precision, Kappa coefficient (Kappa), model parameters, and floating point operations (FLOPs). The mIoU is used to measure the segmentation accuracy of the model on all categories and the mPA is used to measure the classification accuracy on each pixel, both of which are formulated as in Eqs. 12–13.

$$mIoU = \frac{1}{k+1}\sum_{i=0}^{k}\frac{p_{ii}}{\sum_{j=0}^{k}p_{ij}+\sum_{j=0}^{k}p_{ji}-p_{ii}} \tag{12}$$

$$mPA = \frac{1}{k+1}\sum_{i=0}^{k}\frac{p_{ii}}{\sum_{j=0}^{k}p_{ij}} \tag{13}$$

where k represents the number of object categories, with $k+1$ accounting for the sum of object and background categories. Here, i represents a positive category. The term p_{ii} refers to the true positives (TP), which is the count of pixels in category i that are correctly categorized. The term p_{ij} represents the false negatives (FN), indicating the count of pixels in category i that are incorrectly categorized as another category j. The term p_{ji} represents the false positives (FP), which is the count of pixels in category j that are mislabeled as category i. Lastly, p_{jj} is the true negatives (TN). Based on these definitions, the formula of the Precision can be as shown in Eq. 14.

$$Precision = \frac{TP}{TP+FP} \tag{14}$$

The Kappa coefficient is used as a measure of classification accuracy to test its consistency and is formulated as shown in Eq. 15 below.

$$Kappa = \frac{P_o - P_e}{1 - P_e} \tag{15}$$

where P_o denotes the proportion of model-predicted labels that agree with the true labels, and P_e denotes the random agreement probability. Within the scope of our analysis, the grading scale is as follows: scores between 0.41 and 0.60 are categorized as moderate agreement, those ranging from 0.61 to 0.80 are considered substantial agreement, and anything from 0.81 to 1.00 indicates an almost perfect agreement.

4.4 Ablation study

To evaluate the performance of each component of the proposed PL-UNet, we conduct a series of ablation experiments on the TTPLA dataset and the results are shown in Table 1. We also discuss the effectiveness of each modification module independently based on the Baseline network framework, and the results are shown in Table 2. It should be noted that the same training configurations as PL-UNet, including the loss function and optimizer, etc., are used in all experiments, which ensures the fairness of the experiments.

The incorporation of EfficientNetV2-S has led to significant enhancements in all performance metrics compared to the Baseline model. When it comes to the model's capabilities, mIoU is improved by 7.57%, mPA is improved by 10.07%, Precision is improved by 0.81%, and the Kappa coefficient is improved by 0.028. In addition, EfficientNetV2-S reduces the model parameters by 32.57% compared to Baseline. This demonstrates that the EfficientNetV2-S is not only able to extract features effectively, but also has the advantage of low computational requirements.

Based on this, we further integrate the MSAG module. Compared with the introduction of **EfficientNetV2-S**, our improvement achieves 1.27% and 1.55% on mIoU and mPA, respectively, and the Precision and Kappa coefficient rise to 84.54 and 0.829, respectively. Based on Baseline, the MSAG module contributes more significantly to the model performance improvement, increasing mIoU and mPA by 3.83% and 4.52%, respectively. Thus the MSAG module is able

Table 1 Ablation study on the TTPLA dataset

Methods					mIoU	mPA	Precision	Kappa	Params
Baseline	Efficient-NetV2-S	MSAG	DWF	CPPB	(%)	(%)	(%)		(M)
✓					69.33	77.96	83.64	0.796	31.04
✓	✓				76.90	88.03	84.45	0.824	20.93
✓	✓	✓			78.17	89.58	84.54	0.829	20.80
✓	✓	✓	✓		79.12	89.66	85.75	0.841	21.38
✓	✓	✓	✓	✓	79.98	90.07	86.40	0.842	21.27

Table 2 Analysis of the effectiveness of modular improvement based on baseline

Methods	mIoU (%)	mPA (%)	Precision (%)	Kappa	Params (M)
Baseline	69.33	77.96	83.64	0.796	31.04
Baseline + EfficientNetV2-S	76.90	88.03	84.45	0.824	20.93
Baseline + MSAG	73.16	82.48	84.74	0.819	31.16
Baseline + DWF	73.88	83.62	84.43	0.823	32.44
Baseline + CPPB	72.22	81.44	84.28	0.811	35.92

to capture important features more accurately and reduce irrelevant information, proving its effectiveness.

In addition, the DWF module significantly improves performance by optimizing Baseline's jump connections and up-sampling feature fusion. Specifically, mIoU is improved by 4.55% and mPA by 5.66%. These improvements demonstrate the effectiveness of the DWF module in dynamic feature integration. Based on **EfficientNetV2-S + MSAG**, the DWF module further optimizes the model performance, resulting in a 0.95% improvement in mIoU and a 0.08% increase in mPA, as well as an increase in Precision and Kappa coefficient to 85.75 and 0.841, respectively. due to the flexibility of the DWF structure, the model is able to better retain key details in complex aerial power line scene can better retain key details and reduce background interference.

Further, the CPPB module optimizes the two 3×3 standard convolutional layers in Baseline. This improvement is demonstrated on the Baseline model as a 2.89% boost in mIoU and a 3.48% increase in mPA. This demonstrates the effectiveness of the CPPB module in improving model recognition accuracy. Based on **EfficientNetV2-S + MSAG + DWF**, the introduction of the CPPB module further improves the mIoU by 0.86%, the mPA by 0.41%. The Precision and Kappa coefficient increased to 86.40 and 0.842, respectively. This module enables the model to be more clearly segmented under complex high-altitude backgrounds through refined feature selection and enhancement.

Compared with Baseline, PL-UNet achieves an overall improvement of 10.65% in mIoU, an increase of 12.11% in mPA, an increase in Precision from 83.64 to 86.40, and an increase in Kappa coefficient from 0.796 to 0.842. It fully demonstrates its efficiency and robustness in aviation power line scenarios. In addition, by comparing the training time, PL-UNet shows a slight reduction in training time compared to Baseline, by about 15 min. Although the reduction in time is relatively small, this result suggests that PL-UNet can provide some improvement in training efficiency while maintaining high performance. Meanwhile, the model parameters are reduced by 31.48%, which not only improves the detection speed but also optimizes the use of computational resources. These improvements show that PL-UNet is able to significantly improve resource efficiency while maintaining high detection accuracy, providing a better solution for aerial power line inspection in practical applications.

4.5 Impact analysis of loss functions

Within our proposed network for power line segmentation, the selection of the loss function profoundly influences the model's efficacy. To elucidate the impact of various loss functions, we performed comparative experiments employing a range of loss functions, with the outcomes presented in Table 3. Notably, the model's accuracy was markedly enhanced by incorporating the Dice loss, surpassing the results obtained with solely weighted cross-entropy loss. This highlights the Dice loss's effectiveness in addressing class imbalance and enhancing network performance.

4.6 Comparison with advanced methods

To substantiate the efficacy of PL-UNet, we have conducted a performance comparison with several contemporary semantic segmentation models introduced in recent years, with the outcomes depicted in Table 4.

The results show that PL-UNet significantly outperforms the other networks in terms of performance. Specifically, the mIoU of PL-UNet is 14.01%, 7.70%, 5.15% and 19.63% higher than that of PSPNet, Deeplabv3+, HRNet and DANet, respectively. Relative to the SegFormer, which is among the more advanced models developed recently,

Table 3 Experimental analysis of loss function performance

Model	Loss function	mIoU (%)	mPA (%)	Precision (%)	Kappa
PL-UNet	CE loss	74.23	88.09	80.80	0.796
PL-UNet	CE loss + Dice loss	79.98	90.07	86.40	0.842

Table 4 Comparative experiments of advanced methods

Model	Backbone	mIoU(%)	Parameters(M)	Speed/fps
U-Net [16]	×	69.33	31.04	44.46
PSPNet [36]	ResNet50	65.97	46.71	**85.15**
Deeplabv3+ [37]	Xception	72.28	54.71	61.23
DANet [38]	ResNet101	60.35	66.43	83.26
HRNet [39]	×	74.83	29.54	41.12
SegFormer [40]	×	77.53	27.35	51.28
PL-Deeplab [13]	MCNet	78.20	25.44	45.00
PL-UNet(Ours)	EfficientNetV2-S	**79.98**	**21.27**	56.86

PL-UNet exhibits an mIoU improvement of 2.45%. Compared to PL-Deeplab, the latest network model about power line segmentation, the mIoU of the PL-UNet is 1.78% higher. In addition, when the average inference speed is tested by the pictures of the test set, PL-UNet is about 28 fps slower compared to PSPNet and about 26 fps slower compared to DANet. It is also about 4 fps slower compared to Deeplabv3+. On the other hand, PL-UNet is about 16 fps faster compared to HRNet and about 6 fps faster compared to SegFormer. Lastly, PL-Deeplab is about 12 fps faster compared to PL-Deeplab. While PSPNet and DANet perform well in terms of inference speed, they perform poorly in terms of mIoU. And among several networks with relatively close mIoU, PL-UNet achieves higher fps. This shows that PL-UNet not only improves significantly in segmentation accuracy but also exhibits higher efficiency in real-time. PL-UNet can better satisfy the dual needs of speed and accuracy in the task of powerline segmentation.

4.7 Qualitative analysis

As shown in Fig. 5, to be able to better demonstrate the segmentation performance of PL-UNet, we selected some sample images from the TTPLA dataset for experiments and compared them with other networks.

In Fig. 5a and d, the performance of each network in power line segmentation varies significantly. PSPNet and DANet have unsatisfactory segmentation results, with severe power line breaks and even complete loss. U-Net, HRNet, Deeplabv3+ and SegFormer have improved the continuity of the power lines, but there are still small breaks. In contrast, PL-Deeplab performed more consistently and was able to segment most of the power lines, but misclassified the poles. PL-UNet performs best in these scenes with no misclassification. And the power line segmentation is complete, and the fine power lines can also be accurately recognized, with more delicate edge processing. In Fig. 5b and e, facing the complex background and tree occlusion, there is also a big difference in the segmentation performance of the networks in terms of power lines. PSPNet and DANet have poor segmentation in the occluded region, and a large number of power lines are omitted to be detected. SegFormer and HRNet have improved the segmentation effect to a certain extent, but the problem of segmentation discontinuity is still more obvious. DeeplabV3+ has discontinuous segmentation with trees in the background. U-Net and PL-Deeplab are able to segment more power lines in the occluded region,

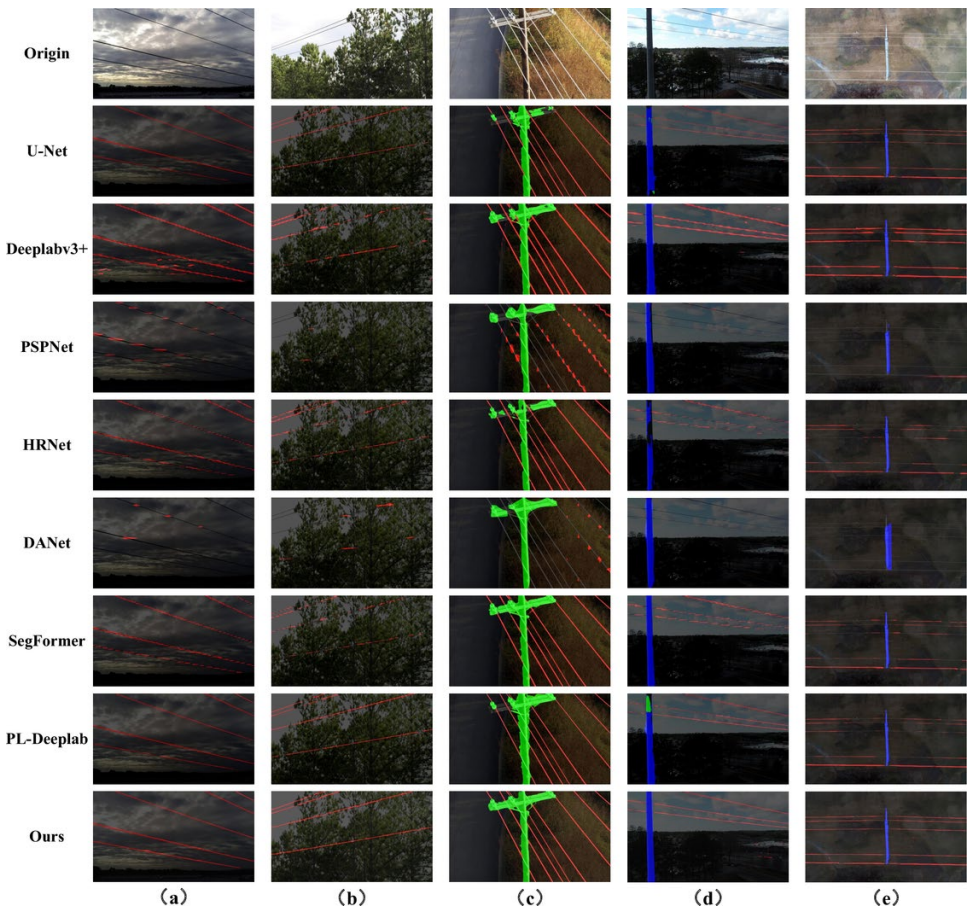

Fig. 5 Segmentation results for different networks on the TTPLA dataset

but still cannot completely avoid the omission of fine lines. And PL-UNet performs well and the power lines are still segmented completely even in the case of tree occlusion and complex background. In Fig. 5c, the segmentation effect of each network in the connection part of the tower and the power line varies. PSPNet and DANet have poor segmentation results, and the connection part of the tower and the power line is fuzzy or even lost. U-Net, HRNet, Deeplabv3+ and PL-Deeplab segmentation results are better, but the processing of the tower is still not fine enough. On the other hand, SegFormer and PL-UNet are more accurate in the segmentation of the connection part of the pole tower and power line, with complete details and clear edge processing. In summary, the segmentation performance of PL-UNet is better than other networks in all kinds of scenes. Whether in the light and dark background, complex background or the connection area of poles and towers, PL-UNet can effectively segment the power lines, especially in the integrity of the fine power lines and the edge detail processing. This indicates that PL-UNet has significant advantages in feature extraction, multi-scale information fusion and detail processing, and delivers a more efficient and dependable approach for the segmentation of power lines.

5 Conclusion

To tackle the challenges of background interference and similar-looking elements in complex aerial power line scenes, as well as the difficulty in distinguishing closely packed power lines, we have proposed a segmentation model. This improved model, based on U-Net and named PL-UNet, addresses these issues effectively. In the encoder, we employ EfficientNetV2-S, which has a compact set of parameters, to expedite feature extraction. In the decoder, Multi-scale Attention Gate are introduced to reduce interference from complex backgrounds. It concentrates on crucial regional features, offering a broader receptive field and lower computational demands. Dynamic Weighted Fusion in the decoder flexibly fuses the features through an adaptive weighting mechanism, which enhances the robustness of feature expression and improves the problem of power line boundary blurring to a certain extent. Cross-stage Partial Pyramid Block enhances the ability to capture fine targets through multi-scale feature extraction and cross-stage feature fusion, providing improvements in the continuity of power line slender structures. The experimental results show that PL-UNet achieves a mIoU of 79.98% on the TTPLA dataset with 21.27M model parameters, 40.91G floating-point operations. The detection speed of 56.86 fps is obtained after testing, which proves the stability and real-time performance of the model. For subsequent studies, a more lightweight network structure can be explored and the robustness under some complex scenes(e.g., dense fog, rain, etc.) can be increased to further improve the model's adaptability in complex environments and efficiency in practical applications.

Data availability The data supporting this study's findings are available from the corresponding author upon reasonable request.

References

1. Santos, Tiago, Cunha, Tiago, Dias, André, Moreira, Antóio Paulo, Almeida, José: UAV Visual and Thermographic Power Line Detection Using Deep Learning. *Sensors*, 24(17):5678, August (2024)
2. Chang, X., Li, Q., Zhou, Q., Zhang, S., Dabing, Yu., Ma, Y.: Power Line-Guided Automatic Electric Transmission Line Inspection System. IEEE Trans. Instrum. Meas. **71**, 1–18 (2022)
3. Azevedo, F., Dias, A., Almeida, J., Oliveira, A., Ferreira, A., Santos, T., Martins, A., Silva, E.: LiDAR-Based Real-Time Detection and Modeling of Power Lines for Unmanned Aerial Vehicles. Sensors **19**(8), 1812 (2019)
4. Senthilnath, J., Kumar, Abhishek, Jain, Anurag, Harikumar, K., Thapa, Meenakumari, Suresh, S., Anand, Gautham, Benediktsson, Jon Atli: BS-McL: Bilevel Segmentation Framework With Metacognitive Learning for Detection of the Power Lines in UAV Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–12, (2022)
5. Zhao, W., Dong, Q., Zuo, Z.: A Method Combining Line Detection and Semantic Segmentation for Power Line Extraction from Unmanned Aerial Vehicle Images. Remote Sensing **14**(6), 1367 (2022)
6. Wang, Lei, Huang, Wei, Chen, Yiping, Chen, Hao, Liu, Chuang, Yin, Jun: Complex Substation Scene Segmentation Method Based on Optimized Straight Line Detection. In *2023 China Automation Congress (CAC)*, pages 4616–4621, Chongqing, China, November (2023). IEEE
7. Zhao, L., Yao, H., Fan, Y., Ma, H., Li, Z., Tian, M.: Power Line Detection for Aerial Images Using Object-Based Markov Random Field With Discrete Multineighborhood System. IEEE Geosci. Remote Sens. Lett. **21**, 1–5 (2024)
8. Gao, Z., Yang, G., Li, E., Liang, Z., Guo, R.: Efficient Parallel Branch Network With Multi-Scale Feature Fusion for Real-Time Overhead Power Line Segmentation. IEEE Sens. J. **21**(10), 12220–12227 (2021)
9. Yang, L., Fan, J., Shuai, X., Li, E., Liu, Y.: Vision-Based Power Line Segmentation With an Attention Fusion Network. IEEE Sens. J. **22**(8), 8196–8205 (2022)
10. Choi, Hyeyeon, Yun, Jong Pil, Kim, Bum Jun, Jang, Hyeonah, Kim, Sang Woo: Attention-Based Multimodal Image Feature Fusion Module for Transmission Line Detection. *IEEE Transactions on Industrial Informatics*, 18(11):7686–7695, November (2022)
11. Yang, L., Fan, J., Huo, B., Li, E., Liu, Y.: PLE-Net: Automatic power line extraction method using deep learning from aerial images. Expert Syst. Appl. **198**, 116771 (2022)
12. Abdelfattah, R., Wang, X., Wang, S.: PLGAN: Generative Adversarial Networks for Power-Line Segmentation in Aerial Images. IEEE Trans. Image Process. **32**, 6248–6259 (2023)
13. Chen, G., Hao, K., Wang, B., Li, Z., Zhao, X.: A power line segmentation model in aerial images based on an efficient multi-branch concatenation network. Expert Syst. Appl. **228**, 120359 (2023)

14. Abdelfattah, Rabab, Wang, Xiaofeng, Wang, Song: TTPLA: An Aerial-Image Dataset for Detection and Segmentation of Transmission Towers and Power Lines. In Hiroshi Ishikawa, Cheng-Lin Liu, Tomas Pajdla, and Jianbo Shi, editors, *Computer Vision–ACCV 2020*, volume 12627, pages 601–618. Springer International Publishing, Cham, (2021). Series Title: Lecture Notes in Computer Science
15. Phan, Trong Huy, Yamamoto, Kazuma: Resolving Class Imbalance in Object Detection with Weighted Cross Entropy Losses. (2020)
16. Ronneberger, Olaf, Fischer, Philipp, Brox, Thomas. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*, volume 9351, pages 234–241. Springer International Publishing, Cham, (2015). Series Title: Lecture Notes in Computer Science
17. Simonyan, Karen, Zisserman, Andrew.: Very Deep Convolutional Networks for Large-Scale Image Recognition, April (2015). arXiv:1409.1556 [cs]
18. He, Kaiming, Zhang, Xiangyu, Ren, Shaoqing, Sun, Jian: Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, Las Vegas, NV, USA, June (2016). IEEE
19. Chollet, François: (2017) Xception: Deep Learning with Depthwise Separable Convolutions, April. arXiv:1610.02357 [cs]
20. Szegedy, Christian, Liu, Wei, Jia, Yangqing, Sermanet, Pierre, Reed, Scott, Anguelov, Dragomir, Dumitru Erhan, Vanhoucke, Vincent, and Rabinovich, Andrew. Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–9, Boston, MA, USA, June (2015). IEEE
21. Howard, Andrew G., Zhu, Menglong, Bo Chen, Kalenichenko, Dmitry, Wang, Weijun, Weyand, Tobias, Andreetto, Marco, Adam, Hartwig: MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, April (2017). arXiv:1704.04861 [cs]
22. Tan, Mingxing, Le, Quoc V.: EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, September (2020). arXiv:1905.11946 [cs]
23. Han, Kai, Wang, Yunhe, Qi Tian, Guo, Jianyuan, Xu, Chunjing, Xu, Chang.: GhostNet: More Features from Cheap Operations, March (2020). arXiv:1911.11907 [cs]
24. Tan, Mingxing, Le, Quoc V.: EfficientNetV2: Smaller Models and Faster Training, June (2021). arXiv:2104.00298 [cs]
25. Wang, Chien-Yao, Liao, Hong-Yuan Mark, Yeh, I.-Hau, Wu, Yueh-Hua, Chen, Ping-Yang, and Hsieh, Jun-Wei: CSPNet: A New Backbone that can Enhance Learning Capability of CNN, November (2019). arXiv:1911.11929 [cs]
26. Wang, Chien-Yao, Bochkovskiy, Alexey, Liao, Hong-Yuan Mark: YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, July (2022). arXiv:2207.02696 [cs]
27. Lou, Ange, Guan, Shuyue, Loew, Murray: CaraNet: Context Axial Reverse Attention Network for Segmentation of Small Medical Objects
28. Liu, Yang, Gou, Peng, Nie, Wei, Xu, Nuo, Zhou, Tianyu, Zheng, Yalan, Wang, Peng: A Power Line Detection Method Combining Semantic Segmentation and Object-Based Analysis. In *2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, pages 540–546, Chongqing, China, June (2022). IEEE
29. Liu, Yang, Ouyang, Ninglei, Gou, Peng, Nie, Wei, Liang, Jing (2022) Fast power line detection based on semantic flow. In *2022 IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pages 230–235, Beijing, China, October. IEEE
30. Oktay, Ozan, Jo Schlemper, Folgoc, Loic Le, Lee, Matthew: Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y. Hammerla, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Attention U-Net: Learning Where to Look for the Pancreas, May (2018). arXiv:1804.03999 [cs]
31. Li, X., Qin, X., Huang, C., Yuer, L., Cheng, J., Liansheng Wang, O., Liu, J.S., Yuan, C.: SUnet: A multi-organ segmentation network based on multiple attention. Comput. Biol. Med. **167**, 107596 (2023)
32. Rahman, Md Mostafijur, Munir, Mustafa, Marculescu, Radu: EMCAD: Efficient Multi-scale Convolutional Attention Decoding for Medical Image Segmentation, May (2024). arXiv:2405.06880 [eess]
33. Ramachandran, Prajit, Zoph, Barret, Le, Quoc V.: Searching for Activation Functions, October (2017). arXiv:1710.05941 [cs]
34. Yan-Xue, W., Kai, D., Wang, X.-J., Min, F.: Misclassification-guided loss under the weighted cross-entropy loss framework. Knowl. Inf. Syst. **66**(8), 4685–4720 (2024)
35. Li, Jian, Liu, Kongyu, Hu, Yating, Zhang, Hongchen, Heidari, Ali Asghar, Chen, Huiling, Zhang, Weijiang, Algarni, Abeer D., Elmannai, Hela: Eres-UNet++: Liver CT image segmentation based on high-efficiency channel attention and Res-UNet++. *Computers in Biology and Medicine*, 158:106501, May (2023)
36. Zhao, Hengshuang, Shi, Jianping, Qi, Xiaojuan, Wang, Xiaogang, Jia, Jiaya: Pyramid Scene Parsing Network
37. Chen, Liang-Chieh, Zhu, Yukun, Papandreou, George, Schroff, Florian, Adam, Hartwig: Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation, August (2018). arXiv:1802.02611 [cs]
38. Fu, Jun, Liu, Jing, Tian, Haijie, Li, Yong: Dual Attention Network for Scene Segmentation
39. Wang, Jingdong, Sun, Ke, Cheng, Tianheng, Jiang, Borui, Deng, Chaorui, Zhao, Yang, Liu, Dong, Mu, Yadong, Tan, Mingkui, Wang, Xinggang, Liu, Wenyu, Xiao, Bin: Deep High-Resolution Representation Learning for Visual Recognition, March (2020). arXiv:1908.07919 [cs]
40. Xie, Enze, Wang, Wenhai, Yu, Zhiding, Anandkumar, Anima, Alvarez, Jose M., Luo, Ping: SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers, October (2021). arXiv:2105.15203 [cs]

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).
Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.
These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.
We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.
While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.
These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.
To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.
Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.
If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com