ISSN: 2414-1895 DOI: 10.6919/ICJE.202504_11(4).0023

Transmission Line Defect Detection Method Integrating Attention Mechanism and Inverted Residual Design

Xueming Zhai[a], Chao Ling[b]

North China Electric Power University, Baoding 071003, China

[a]zxm3165@126.com, [b]1643841489@qq.com

Abstract

To address the significant workload and high missed detection rate in the defect detection of transmission line inspection images, this paper proposes a defect detection method for transmission lines that balances real-time performance and accuracy. Based on the YOLOv8 model, the BoNAM module structure is designed to suppress background noise in transmission line images and reduce the dimensionality of input feature maps. The iRMB module is employed to efficiently extract backbone network features, while the SPD-Conv module is utilized to more precisely identify the missing locations of bolts and pins. Multiple ablation experiments were conducted on the improved algorithm components, with the enhanced method achieving an average precision of 86.9%, an improvement of 4.3% compared to the baseline model. Comparative experiments with mainstream object detection algorithms demonstrate that the proposed method outperforms others in terms of both accuracy and speed.

Keywords

Transmission Line; Defect Detection; Attention Mechanism; Feature Fusion; Feature Extraction.

1. Introduction

The transmission and distribution lines are key components of the power system, characterized by their large scale, exposure to complex and variable natural climatic conditions, and high electrical load [1,2]. Therefore, regularly inspecting transmission lines and performing timely safety maintenance are crucial for ensuring the stable operation of the power system. However, since transmission lines are exposed to the natural environment and are subject to harsh weather conditions such as heavy rain and prolonged sun exposure, critical components of the transmission lines are prone to safety hazards. These hazards may lead to line failures, posing significant threats to people's daily lives and the overall safety of electricity usage in society [3,4].

In the early days, transmission line inspections were primarily conducted manually. However, this method had significant drawbacks, including high labor intensity and low inspection efficiency [5,6]. Prolonged fatigue among workers often led to a substantial increase in missed and erroneous detections. Furthermore, during natural disasters such as floods, earthquakes, and landslides, manual inspections became even more difficult, making power inspection results unpredictable. Although drone-based inspections were later widely adopted and applied, the vast number of collected images still required manual verification, resulting in no significant improvement in inspection efficiency.

With the development of deep learning and object detection technology, intelligent power grid inspection has garnered widespread attention [7,8]. The vast amount of image data collected by drones can be processed using object detection technology to replace human visual inspection, enabling fast and accurate identification of defect locations and types.Existing deep learning-based object detection technologies can be categorized into two major types based on whether they require the prior generation of candidate boxes. Algorithms that require candidate box generation are known as two-stage object detection algorithms, with the R-CNN [9] series being the most representative. Based on the R-CNN model, various object detection algorithms have been developed, including SPP-Net [10], Fast R-CNN [11], and Faster R-CNN [12]. Since 2014, the R-CNN series has rapidly evolved, achieving high accuracy in various downstream tasks. However, due to its large number of parameters, slow inference speed, and the need for candidate box generation, this model is difficult to apply in real-time scenarios such as video streaming.In contrast, single-stage object detection algorithms eliminate the candidate box generation step. After extracting features using a CNN network, they directly perform object classification and bounding box regression on the feature map. The most representative algorithms in this category are SSD (Single Shot MultiBox Detector) [13] and the YOLO (You Only Look Once) series. These algorithms achieve a better balance between inference speed and accuracy, making them well-suited for applications in intelligent driving, video surveillance, and other fields. In the field of object detection, single-stage algorithms have largely replaced two-stage detection methods.

Many researchers have conducted extensive studies on this topic.Literature [14] proposed the Box-Point Detector method, which consists of a deep convolutional neural network and two parallel convolutional heads. This method preserves more information for small-sized faults, effectively improving the detection accuracy of minor insulator defects. However, it struggles to adapt to small defect detection in complex background interference scenarios.Literature [15] introduced RepYOLO based on the YOLOv5 network, integrating RepVGG, Diversified Branch Block (DBB), Efficient Channel Attention (ECA), and an improved Spatial Pyramid Pooling (SPP) module. This approach improved accuracy by 1.2% compared to YOLOv5 while tripling inference speed.Literature [16] built upon YOLOv7 by designing a Dual-branch Serial Attention mechanism and a Well-connected Feature Pyramid Network (WCFPN). These enhancements increased the weight of defect regions and facilitated the full fusion of semantic information at different levels, leading to strong performance in insulator defect detection.Literature [17] addressed challenges such as complex backgrounds and small defect targets in transmission line images. By optimizing the energy function and employing the Normalized Wasserstein Distance (NWD) loss function based on the YOLOv5s model, the study achieved excellent results in detecting small defects in bird nests and insulators.Literature [18] utilized U-Net as the base network, incorporating Depthwise Separable Convolution and the ECA attention mechanism. These modifications highlighted critical network features while reducing the number of network parameters, resulting in improved performance in detecting defects in transmission conductors.

In summary, this paper addresses challenges in transmission line defect detection, such as severe background interference and small defect sizes, by proposing an improved YOLOv8-based detection method. The main contributions are as follows:First, a Normalization-based Attention Module (NAM) [19] is introduced. Inspired by the Bottleneck structure in the C3 module of YOLOv5, the BoNAM module is designed to maintain a lightweight structure while suppressing complex background weights to achieve higher accuracy.Next, the backbone network incorporates the Inverted Residual Mobile Block (iRMB) [20], which combines the efficiency of convolutional neural networks with the ability of Transformers [21] to capture long-range contextual dependencies.Finally, an SPD-Conv module, combining the Focus module and Conv module, is introduced into the backbone network for downsampling feature maps. This preserves more fine-grained feature information, enhancing the detection of small defect targets.

Additionally, to address the issue of limited defect variety in existing overhead transmission line datasets, this paper collects 1927 images of overhead transmission line defects from the internet and through direct photography. The images are annotated using the LabelImg tool in YOLO format, resulting in the creation of a small-scale overhead transmission line defect dataset.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202504_11(4).0023

2. Related Knowledge

2.1 Data Augmentation

Deep learning-based object detection techniques often require large amounts of high-quality images and annotations. Acquiring large-scale datasets is expensive and challenging. Data augmentation techniques generate additional training samples by transforming original images, enhancing model generalization and robustness. YOLOv8 primarily employs Mosaic[22] and Mixup[23] for data augmentation. Mosaic combines four different images into one, adjusting positions and sizes randomly. Mixup blends two images using linear interpolation, creating a new sample with weighted pixel values.

2.2 Attention Mechanism

In traditional convolutional neural networks, all input features are treated equally, limiting their ability to emphasize critical information. Attention mechanisms assign different weights to feature information, helping models focus on defect regions. SE (Squeeze-and-Excitation)[24] enhances feature representation by learning channel-wise attention. CBAM (Convolutional Block Attention Module) [25] operates on both channel and spatial dimensions. The channel attention mechanism learns attention weights through global average pooling and fully connected layers, while the spatial attention module learns spatial attention weights by stacking convolution kernels of different scales.

2.3 YOLOv8 Algorithm

YOLOv8, developed by Ultralytics, builds upon YOLOv5 with continuous optimizations. It consists of three main parts: a feature extraction backbone, a feature fusion neck, and classification/regression detection heads. YOLOv8 replaces the C3 module in YOLOv5 with the gradient-rich C2f module in the backbone. The PAN structure in the neck fuses multi-scale features bidirectionally. The detection head separates classification and regression tasks and transitions from anchor-based to anchor-free detection, improving adaptability.

3. Transmission Line Defect Detection Method Integrating Attention Mechanism and Inverted Residual Designs

This paper proposes an improved YOLOv8-based transmission line defect detection method, as shown in Figure 1. The method enhances the backbone network by integrating the BoNAM attention mechanism and iRMB module to improve feature extraction capability, while the SPD-Conv module enhances small-object detection performance.

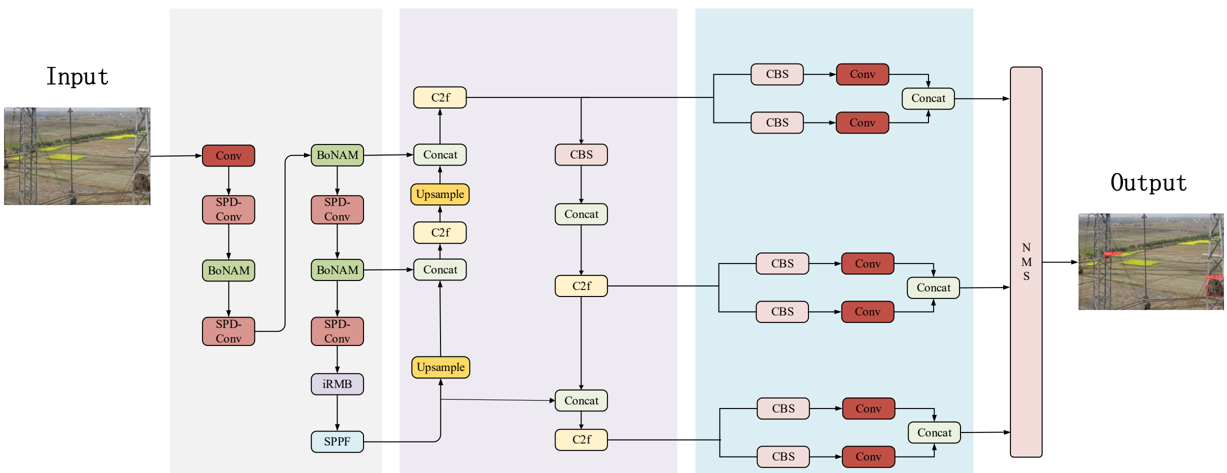

Figure 1. Algorithm Structure

ISSN: 2414-1895
DOI: 10.6919/ICJE.202504_11(4).0023

3.1 Design of BoNAM Attention Mechanism

Transmission lines are mostly located in outdoor environments with complex and variable background information, making traditional object detection methods highly susceptible to background interference. To address this issue, this paper introduces a Normalization-based Attention Module (NAM) and designs the BoNAM attention mechanism based on the Bottleneck structure. This module is integrated into the backbone network to suppress background features in transmission line images and reduce the dimensionality of feature maps, thereby enhancing the feature extraction capability of the backbone network without introducing a large number of additional parameters, as shown in Figure 2.

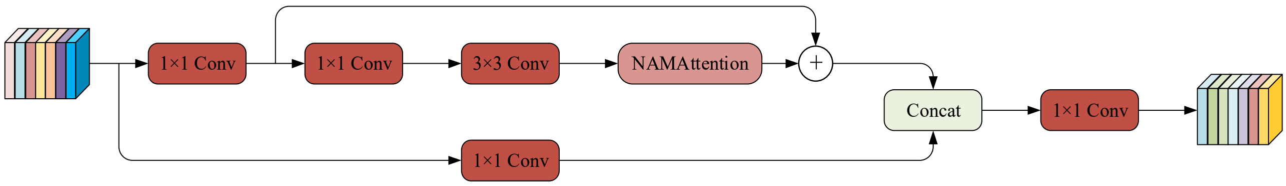

Figure 2. BoNAM Attention Mechanism

The input to BoNAM comes from the feature map of the previous network layer and is processed through two branches, A and B. The A branch first passes through two 1×1 convolution layers and a 3×3 convolution layer, followed by the NAM attention mechanism. A residual structure is also incorporated to improve the generalization ability of the network. The B branch undergoes a 1×1 convolution before being concatenated with the output of the A branch, enriching the gradient flow during backpropagation. The concatenated result is then processed using a 1×1 convolution layer to enhance the correlation between feature map channels. BoNAM employs multiple 1×1 convolutions to ensure lightweight characteristics while leveraging the efficient NAM attention mechanism to suppress unnecessary weights, ultimately improving the overall model accuracy.

The NAM module introduced in BoNAM provides an efficient and lightweight attention mechanism. NAM is inspired by CBAM and focuses on both spatial and channel attention while eliminating fully connected layers and convolution layers, thereby reducing parameter size and computational cost. This module uses the standard deviation of batch normalization factors to represent attention weights, ensuring its lightweight and efficient nature. This paper adopts the channel attention submodule, whose structure is shown in Figure 3. The input feature map first undergoes batch normalization and is then multiplied by the corresponding attention weights, with the computation process formulated in Equation (1).

$$O = sigmoid\left(W_\gamma\left(BN(F)\right)\right) \tag{1}$$

Where O represents the output feature map, sigmoid denotes the activation function, BN refers to the batch normalization operation, and γ represents the scaling factor for each channel. The calculation of W_γ is shown in Equation (2).

$$W_\gamma = \frac{\gamma_i}{\sum_{j=0}\gamma_j} \tag{2}$$

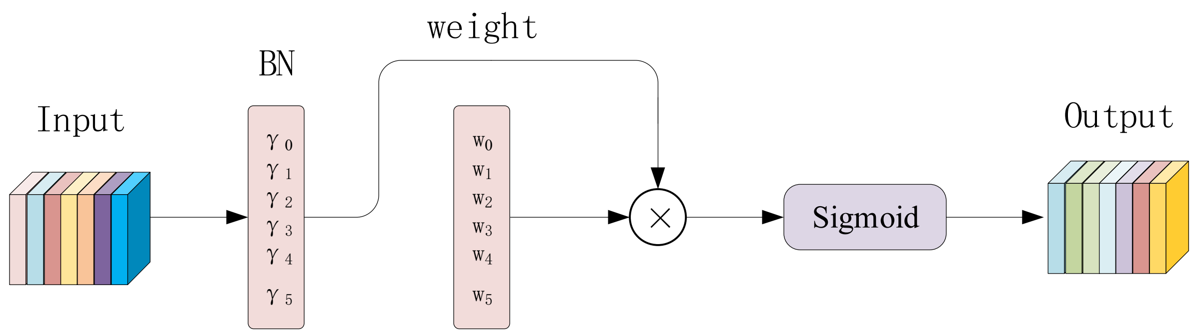

Figure 3. Channel Attention Submodule

3.2 Introduction of SPD-Conv Module

In transmission line defect detection images, the transmission line itself occupies a relatively small portion of the entire image. Within the transmission line, the defect region is even smaller in proportion to the line itself. To accommodate small defect detection in transmission lines, this paper designs the SPD-Conv module, inspired by the Focus module in YOLOv5. The workflow of this module is illustrated in Figure 4.

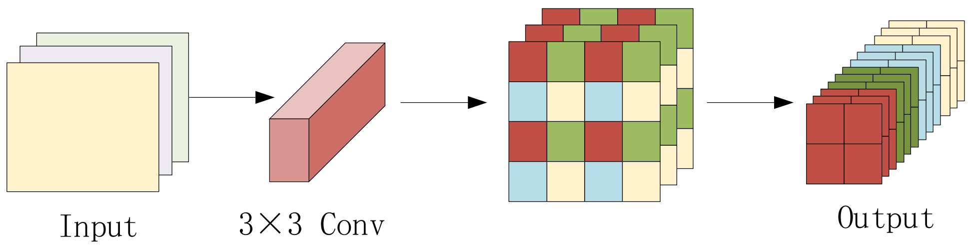

Figure 4. SPD-Conv module

The SPD-Conv module is located in the backbone network, adjacent to the BoNAM module. First, the input feature map undergoes a 3×3 convolution. Then, it is divided into four sub-images based on a 2×2 grid. These four sub-images are then stacked along the channel dimension, forming a new feature map with reduced resolution but an increased number of channels. SPD-Conv utilizes a slicing operation to reduce the feature map size to one-quarter of its original dimensions. This approach preserves more fine-grained image details without introducing significant computational overhead, enabling the backbone network to better extract features of small defect targets.

3.3 Introduction of Inverted Residual Mobile Block

In the penultimate layer of the backbone network, this paper introduces the Inverted Residual Mobile Block (iRMB), which combines convolutional structures with a self-attention mechanism to enhance performance. While this module sacrifices some detection efficiency, it significantly improves detection accuracy. The module structure is shown in Figure 5.

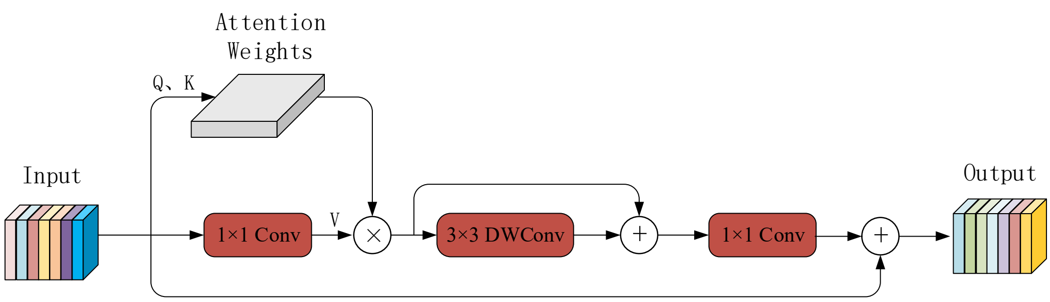

Figure 5. Inverted Residual Mobile Block

ISSN: 2414-1895 DOI: 10.6919/ICJE.202504_11(4).0023

The iRMB module integrates convolutional structures with a self-attention mechanism, allowing it to efficiently capture both local features and long-range contextual dependencies. Its computational process is formulated in Equation (3).

$$O = Conv\big(MHSA(X)\big) \tag{3}$$

Here, Conv represents the convolution operation, MHSA denotes the multi-head self-attention mechanism, X is the input feature map, and O is the output feature map. In this module, the Q (query), K (key), and V (value) matrices are first computed from the input feature map. The Q and K matrices undergo a dot product operation to obtain attention weights, which are then multiplied by the V matrix after passing through a 1×1 convolution. This is followed by a 3×3 depth-wise convolution (DW-Conv), a 1×1 convolution, and a residual connection, as illustrated in Figure 5, to form the final output feature map.The convolutional structures adjacent to the input and output perform channel expansion and reduction, respectively, forming an inverted residual structure. This design significantly reduces the module's parameter count and computational complexity.

4. Experimental Results and Analysis

4.1 Experimental Platform and Parameter Settings

This experiment was conducted on a Windows 10 operating system with an Intel(R) Core(TM) i7-9700 CPU @ 3.00GHz, 32GB of RAM, and an NVIDIA GeForce RTX 2080Ti GPU with 11GB of VRAM. The software environment included Python 3.8.17, the deep learning framework PyTorch 2.0.1, CUDA 11.8, and cuDNN 8.7.0.

During the data preprocessing stage, Mosaic augmentation was applied for data enhancement. The batch size was set to 16, and the input image size was 640×640. The optimizer used was AdamW, with an initial learning rate of 0.01, a learning rate momentum of 0.937, and a weight decay of 0.0005. A total of 300 training epochs were conducted, with a warm-up phase in the first 3 epochs and data augmentation disabled in the final 10 epochs.

4.2 Dataset

A total of 1,927 defect images of overhead transmission lines were collected for this experiment, covering three types of defects: nest obstructions (Nest), damaged insulators (Damaged Insulator), and missing bolt pins (Missing Pin). Among them, 612 images contained nest obstructions, 710 images showed damaged insulators, and 605 images had missing bolt pins.

The dataset was annotated using the LabelImg tool in YOLO format. In comparative experiments, since different algorithms support different annotation formats, the YOLO format was converted to VOC format. The dataset was split into a training set and a test set in an 8:2 ratio, with 1,541 images in the training set and 386 images in the test set.

4.3 Experimental Performance Evaluation Metric

This experiment uses Average Precision (AP) and mean Average Precision (mAP) to evaluate the accuracy of the algorithm, uses the number of parameters to assess the model size, and uses Frames Per Second (FPS) to evaluate the inference speed of the model.

4.4 Ablation Experiments and Analysis

To validate the effectiveness of the module, this paper designs 8 sets of ablation experiments, each training for 300 epochs. The experimental results are shown in Table 1. The baseline model is YOLOv8n. The comprehensive evaluation metric, Average Precision (AP), is used to assess the model's accuracy, where AP1, AP2, and AP3 represent the average precision of the categories of bird's nest foreign objects, insulator damage, and missing bolt pins, respectively. mAP represents the mean average precision of these four categories.

Volume 11 Issue 4, 2025
DOI: 10.6919/ICJE.202504_11(4).0023

Table 1. Ablation experiment results

BoNAM	SPD-Conv	iRMB	MPAN	mAP/%	AP1/%	AP2/%	AP3/%	FPS	Parameters/M
				82.6	83.4	81.7	82.8	**63**	3.0
√				83.6	83.2	82.3	85.4	57	**2.74**
	√			83.6	81.5	83.8	85.4	61	3.27
		√		84.3	87.0	81.7	84.1	62	2.81
			√	83.4	85.6	80.4	84.1	56	3.08
√	√			84.1	84.2	82.1	86.0	56	3.04
√	√	√		85.9	87.0	84.1	86.5	54	3.26
√	√	√	√	86.9	91.5	82.5	86.7	53	3.34

In terms of accuracy, as shown in Table 1, the original YOLOv8n algorithm achieves an mAP of 82.6% on the transmission line defect dataset. The application of the BoNAM method improves the mAP by 1.0%, the SPD-Conv method improves it by 1.0%, and the iRMB method improves it by 1.7%. After applying all the improved methods, the mAP increases to 86.8%, which is a 4.0% improvement over the original algorithm. In terms of speed, the FPS with all the improved algorithms decreases by 20% compared to the original algorithm, but the rate of 74 frames per second still meets the real-time detection requirements. In terms of model size, the number of parameters with all the improved algorithms increases by 11.3% compared to the original algorithm.

4.5 Comparison Experiments with Different Algorithms

To fully validate the effectiveness of the algorithm in this paper, we conduct comparative experiments between the current mainstream and classic object detection algorithms and the algorithm presented in this paper. The experimental results are shown in Table 2. Here, AP1, AP2, and AP3 represent the average precision for the categories of bird's nest foreign objects, insulator damage, and missing bolt pins, respectively. mAP represents the mean average precision of these three categories, FPS assesses the model's speed, and Parameters evaluates the model size.

Table 2. Performance comparison of different algorithms

Method	mAP/%	AP1/%	AP2/%	AP3/%	FPS	Parameters/M
Faster-RCNN	75.7	73.8	76.2	76.3	24	136.7
CenterNet[26]	80.9	79.0	82.1	80.8	22	16.4
SSD	75.3	72.4	82.6	71.7	49	26.3
YOLOv3	76.3	74.9	72.7	81.0	30	103.7
YOLOv5	81.9	83.0	78.6	83.3	45	7.1
YOLOv6	80.1	83.4	77.9	80.1	52	4.2
YOLOv7	82.6	83.7	80.7	82.1	53	6.0
YOLOX	81.6	78.6	83.4	81.8	48	54.1
YOLOv8n	82.3	83.5	81.6	82.8	63	3.0
Ours	86.9	91.5	82.5	86.7	53	3.3

ISSN: 2414-1895 DOI: 10.6919/ICJE.202504_11(4).0023

As shown in Table 2, the algorithm in this paper achieves the highest mAP, with detection speed second only to YOLOv8n. Compared to the Faster-RCNN, CenterNet, and YOLOv3 algorithms, the algorithm in this paper significantly improves both detection accuracy and speed, with mAP increasing by 11.2%, 6.0%, and 10.6%, respectively. Compared to SSD, YOLOv5, YOLOv6, YOLOv7, and YOLOX algorithms, the algorithm in this paper achieves a substantial improvement in detection accuracy with detection speed remaining roughly the same or slightly improved. The mAP increases by 11.6%, 5.0%, 6.8%, 4.3%, and 5.3%, respectively. Notably, the bird's nest foreign object and insulator damage categories show significant improvements, while the missing bolt pin category shows improvements in comparison to other algorithms, with only a 0.9% decrease compared to the YOLOX algorithm. Compared to the baseline model YOLOv8n, the algorithm in this paper introduces more parameters and computation, leading to a slight decrease in detection speed, but it still holds an advantage over other algorithms. Therefore, the algorithm in this paper is more suitable for deployment on edge devices with limited computational resources compared to other algorithms.

4.6 Visualization Results Analysis

This paper provides a visual presentation of the experimental results, as shown in Figure 6, comparing the performance of the YOLOv8n algorithm and the algorithm in this paper for detecting four types of defects in transmission lines. The algorithm in this paper is able to better extract the features of small targets from complex backgrounds, eliminate interference from similar objects and occlusions, and effectively reduce the false detection rate.

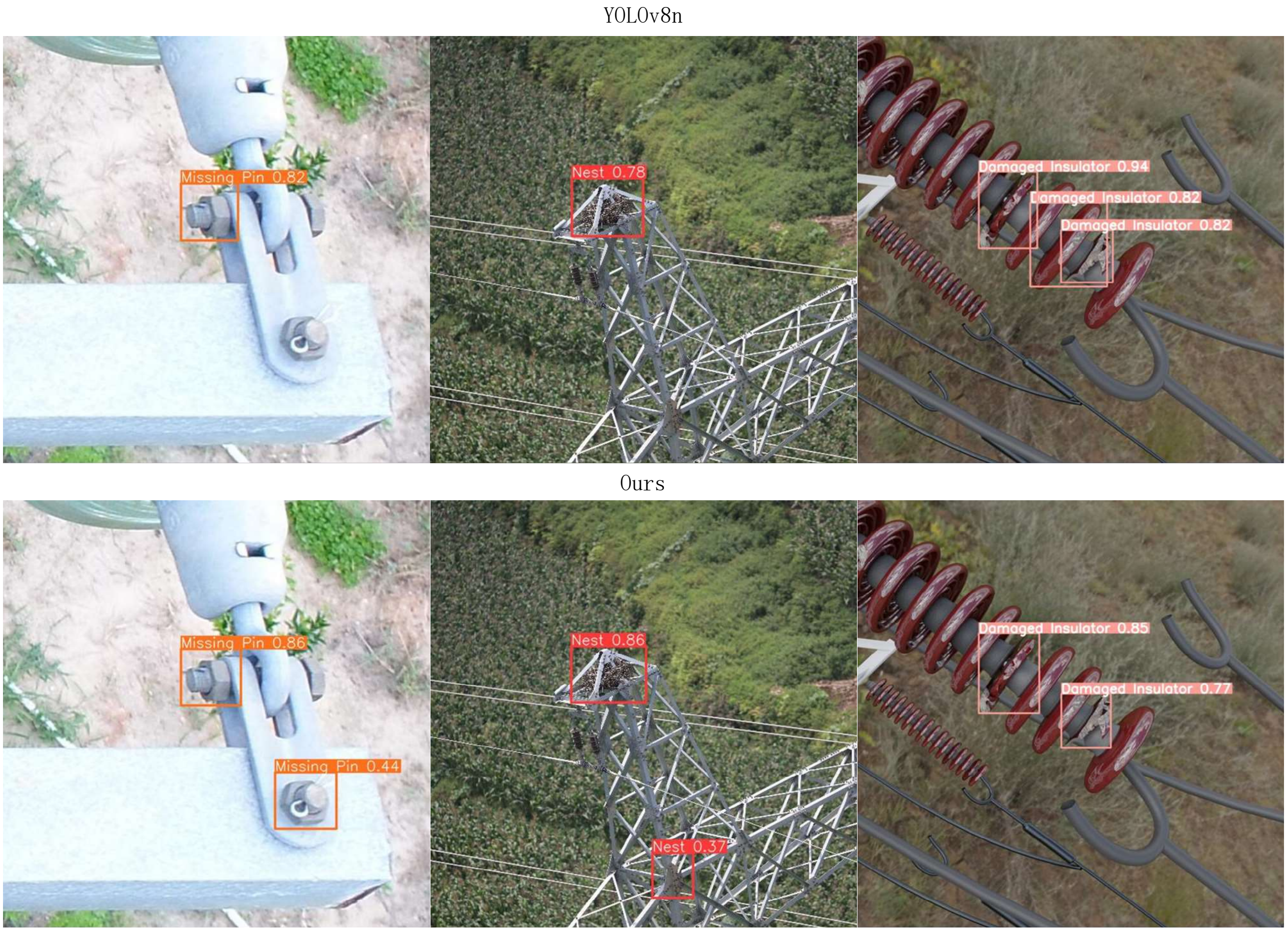

Figure 6. Model comparison visualization

5. Conclusion

This paper proposes an overhead transmission line defect detection method based on the YOLOv8 algorithm. To address challenges such as background interference and small defect targets in transmission lines, the detection performance is effectively improved by optimizing the feature fusion network and constructing new modules. The main contributions of this paper are as follows: the BoNAM module is constructed to suppress the defect background's weight and reduce the dimensionality of the feature maps; the iRMB module is introduced to significantly improve detection accuracy without introducing a large number of parameters and computations; the SPD-Conv module adjusts the three dimensions of the feature map, optimizing the detection of missing bolt pin locations. Experimental results show that the proposed method can effectively handle issues such as severe background interference and large differences in defect targets in transmission lines. Future research could further optimize accuracy and real-time performance while improving the algorithm's robustness in different weather conditions.

References

[1] SUI Yu,NING Pingfan,NIU Pingjuan,et al.Review on mounted UAV for transmission line inspection[J]. Power System Technology,2021,45(9):3636-3648.

[2] MA Fuqi,WANG Bo,DONG Xuzhu,et al.Receptive field vision edge intelligent recognition for ice thickness identification of transmission line[J].Power System Technology,2021,45(6):2161-2169.

[3] LIU Zhiying,MIAO Xiren,CHEN Jing,et al.Review of visible image intelligent processing for transmission line inspection[J].Power System Technology,2020,44(3):1057-1069.

[4] PENG Xiangyang,CHEN Chi,RAO Zhangquan, et al.Safety inspection and intelligent diagnosis of transmission line based on unmanned helicopter of multi sensor data acquisition[J].High Voltage Engineering,2015,1(1):159-166.

[5] PENG Xiangyang,YI Lin,QIAN Jinju,et al.Practical research on patrol inspection system of electric power line based on large scale unmanned helicopter[J].High Voltage Engineering,2020,46(2):384-396.

[6] YU Changle,QU Boyang,ZHU Yuanda,et al.Design of the transmission line inspection system based on UAV[C]//Proceedings of 2020 10th International Conference on Power and Energy Systems (ICPES).Chengdu:IEEE,2020:543-548.

[7] M. M. Hosseini,A. Umunnakwe,M. Parvania,andT. Tasdizen, “Intelligent damage classification andestimation in power distribution poles using unmanned aerial vehicles and convolutional neural networks” IEEE Trans. Smart Grid,vol. 11,no. 4,pp. 3325–3333,Jul. 2020.

[8] Z. Zhao,H. Qi,Y. Qi,K. Zhang,Y. Zhai,and W. Zhao,“Detection method based on automatic visual shape clustering for pin-missing defect in transmission lines,” IEEE Trans. Instrum. Meas.,vol. 69,no. 9,pp. 6080–6091,Sep. 2020.

[9] GIRSHICK R,DONAHUE J,DARRELL T,et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Lake City,2014: 580-587.

[10] He K,Zhang X,Ren S,et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE transactions on pattern analysis and machine intelligence,2015,37(9): 1904-1916.

[11] GIRSHICK R. Fast R-CNN [C]//Proceedings of the IEEE international Conference on Computer Vision. Santiago,2015:1440-1448.

[12] REN S,HE K,GIRSHICK R,et al. Faster R-CNN: towards real-Time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(6): 1149.

[13] LIU W,ANGUELOV D,ERHAN D,et al. SSD: single shot multi-box detector[C]//Proceedings of the 14th European Conference on Computer Vision. Amsterdam the Netherland,2016: 21-37.

[14] Liu X,Miao X,Jiang H,et al. Box-point detector: A diagnosis method for insulator faults in power lines using aerial images and convolutional neural networks[J]. IEEE Transactions on Power Delivery,2021,36(6): 3765-3773.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202504_11(4).0023

[15] Liu M,Li Z,Li Y,et al. A fast and accurate method of power line intelligent inspection based on edge computing[J]. IEEE Transactions on Instrumentation and Measurement,2022,71: 1-12.

[16] Deep Neural Network Model for Transmission Line Defect Detection Based on Dual-Branch Sequential Mixed Mechanism[J]. Computer Science,2024,51(03):135-140.

[17] Ge Zhao,Li Hongwen,Liu Haifeng,et al. A YOLO-GSS-based Real-time Defect Detection Method for Transmission Lines at the Edge End[J/OL]. High Voltage Technology,1-10[2024-12-03].https://doi.org/10.13336/j.1003-6520.hve.20232175.

[18] Zhai Xueming,Li Xiao,Zhai Yujia,et al. Defect Detection of Overhead Transmission Wires Based on Deep Learning [J]. Power System Technology,2023,47(03): 1022-1031.DOI:10.13335/j.1000-3673.pst.2022.0738.

[19] Liu Y,Shao Z,Teng Y,et al. NAM: Normalization-based attention module[J]. arXiv preprint arXiv:2111.12419,2021.

[20] Zhang J,Li X,Li J,et al. Rethinking mobile block for efficient attention-based models[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE Computer Society,2023: 1389-1400.

[21] Vaswani A,Shazeer N,Parmar N,et al. Attention is all you need[J]. Advances in neural information processing systems,2017,30.

[22] BOCHKOVSKIY A,WANG C Y,LIAO H Y M. YOLOv4: optimal speed and accuray of object detection[EB/OL]. [2023-06-29]. https://arxiv.org/abs/2004.10934.pdf.

[23] ZHANG H Y,CISSE M,DAUPHIN Y N,et al. Mixup: beyond empirical risk minimization[EB/OL]. [2023-06-29]. https://arxiv.org/abs/1710.09412.pdf.

[24] HU J,SHEN L,SUN G. Squeeze-and-excitation networks[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press,2018: 7132-7141.

[25] WOO S,PARK J,LEE J Y,et al. CBAM: convolutional block attention module[C]//European Conference on Computer Vision. Cham: Springer,2018: 3-19.

[26] GUO C,FAN B,ZHANG Q,et al. AugFPN: Improving multi-scale feature learning for object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle,2020: 12595-12604.