ISSN: 1813-4890

Improved YOLOv8 Insulator Defect Detection Algorithm

Lifen Li [1,2], Ming Wang [1,2, a,*]

[1] School of Computer Science, North China Electric Power University, Baoding, Hebei, China

[2] Engineering Research Center of the Ministry of Education for Intelligent Computing of Complex Energy Systems, Baoding, Hebei, China

[a] mingwang199902@163.com

Abstract

The role of object detection algorithms in the field of transmission line insulator inspection has become increasingly important. To address the issues of high network complexity and low detection accuracy of insulator defects in complex backgrounds, this paper proposes an improved YOLOv8 algorithm for insulator defect detection based on attention mechanism and multi-scale feature fusion. By incorporating the CBAM attention mechanism at the Backbone, replacing the PANet structure in the Neck with the BiFPN structure, and adopting SIoU as the loss function during the regression process, the YOLOv8 model's feature extraction capability for insulator self-explosion, broken, and flashover regions is enhanced, while mitigating the impact of complex backgrounds such as lighting and vegetation on detection accuracy, thereby improving the model's detection precision. To validate the effectiveness of the proposed algorithm, experiments are conducted on a public dataset of insulators. The improved YOLOv8 model demonstrates excellent performance on real samples, achieving a detection precision of 93.7%, which represents an improvement of 9.4% in P and 2.3% in mAP compared to the original YOLOv8 model, making it suitable for practical inspection tasks.

Keywords

YOLOv8, Detection Algorithm.

1. Introduction

Due to long-term exposure to harsh outdoor environments, insulators are prone to malfunctions, posing significant safety hazards to the safe operation of the transmission network[1]. Therefore, regular attention should be paid to the status of insulators, transmission line inspections should be arranged, and insulator faults should be promptly eliminated. The traditional manual inspection approach faces challenges such as high inspection difficulty, compromised safety assurance, and heavy reliance on human expertise for inspection results, leading to difficulties in ensuring inspection accuracy. With the advancement of technology, unmanned aerial vehicles, robots, and other intelligent inspection methods have gradually replaced manual inspections[2].

Wang et al. [3] proposed using the instance segmentation network Mask R-CNN for insulator defect detection, effectively solving the problem of complex background and small insulator defect objects in drone aerial images, which are difficult to accurately identify. Zhao et al. [4] proposed an enhanced Faster R-CNN model and applied it to locate insulators within complex background images. This method contributes to an improved accuracy in both insulator recognition and fault detection. Qi et al. [5] embedded a dual attention mechanism in Faster R-CNN, effectively avoiding errors and omissions in bolt defect detection of transmission lines. Yang et al. [6] applied the Faster R-CNN model to ground glass density shadow detection in lung CT images and achieved good results, demonstrating the effectiveness of the two-stage object

ISSN: 1813-4890

detection algorithm. However, the two-stage model has too many region proposals and adjacent windows have redundant information, resulting in high computational complexity and slow detection speed. Jia et al. [7] proposed an improved SSD algorithm that introduces Xception deep separable convolution to achieve the detection of small objects on the sea surface. Starting from the YOLOv1 model, the YOLO series of models have iteratively developed versions such as YOLOv2, YOLOv3, and YOLOv4 after years of development. Stefenon et al. [8] introduced a hybrid approach called YOLOu-Quasi-ProtoPNet for the detection and classification of faulty insulators. The proposed method based on DenseNet-161 achieved an F1 score of 0.95165, outperforming similar models in the classification task. Cai et al. [9] introduced the stacked hourglass network and K-means clustering algorithm into

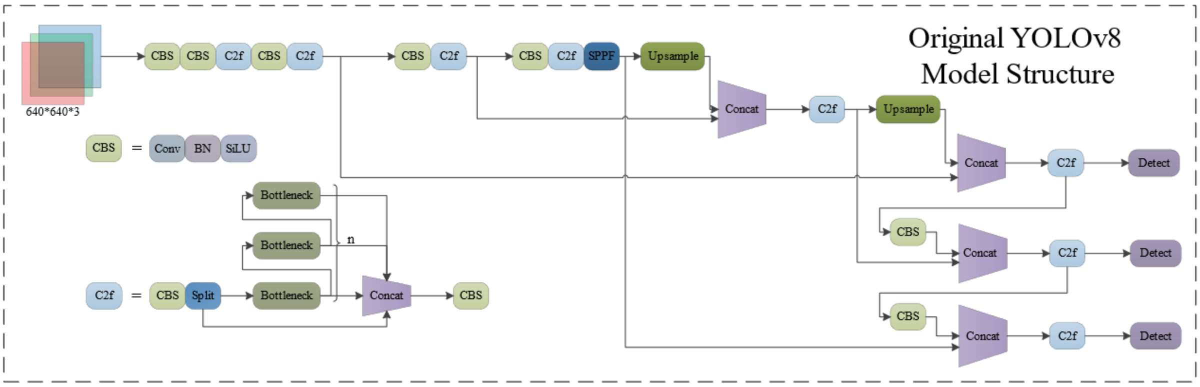

Figure 1 Original YOLOv8 model structure

the YOLOv3 model, which was supplemented by pruning operations to achieve multi person pose evaluation, proving the practical value of the YOLO model. Liu et al.[10] proposed a YOLOv3 model based on SPP (spatial pyramid pooling) network and multi-scale prediction network to detect insulator defects, achieving high detection accuracy. However, the complex network structure and large volume of the model result in slower inference speed. Chen et al. [11] combined the encoding structure of Transformer with the YOLOv5 model to perform pruning operations, reducing the complexity of the model and improving inference speed, but with a decrease in detection accuracy. The aforementioned methods have shown individual improvements in both detection accuracy and inference speed, but they have not managed to achieve a balance between the two. In addition, the method of manual annotation significantly impacts the detection results. Hao et al.[12]proposed a weakly supervised and phased transfer learning method based on YOLOv5 to recognize insulators and different types of ice, such as snow, frost, mixed frost, ice particles, and normal states, which markedly improved the inefficiency associated with manual annotation.

In order to meet the requirements of precise detection of insulator defects on mobile devices such as unmanned aerial vehicles and intelligent inspection robots, and to address the high complexity of the YOLOv5 model and the low detection accuracy for small objects and complex backgrounds, this paper proposes an improved YOLOv8 model. The proposed improved YOLOv8 model based on attention mechanism and multi-scale feature fusion can further improve detection accuracy. Firstly, the attention mechanism of CBAM (Convolutional Block Attention Module) (Woo et al.[13]) is integrated between the fourth C2f module of the YOLOv8 backbone feature extraction network and the SPPF (Spatial Pyramid Pooling Fast) module to enhance the model's feature extraction ability and improve the recognition accuracy of small objects and occluded objects; On this basis, in the feature fusion stage, BiFPN (Bidirectional Feature Pyramid Network) (Tan et al.[14]) structure was used to replace the original PANet (path aggregation network) structure, simplifying the network structure while enhancing the model's feature fusion ability; Finally, taking full account of the impact of the angle loss between the prediction box and the ground truth box on the prediction results in the regression process, SIoU (Gevorgyan[15]) is used as the loss function of the head end regression stage to accelerate the rate of convergence of the model.

2. YOLOv8 Model

The network structure of the YOLOv8 model is illustrated in Figure 1, mainly composed of four parts, including the input end, the backbone feature extraction network, the feature fusion network, and the prediction end. The original YOLOv8 model continues the Mosaic data augmentation technology used by YOLOv5 at the input end, but this method easily makes the model learn some information that is useless for detection. Therefore, YOLOv8 turns off the Mosaic data augmentation technology in the last 10 rounds of training. In the backbone feature extraction network, the YOLOv8 model borrows the design concept of ELAN (Wang et al.[16]) in YOLOv7 and replaces all C3 modules used in YOLOv5 with C2f modules. The structures of C2f modules and C3 modules are shown in Figure 2. The C2f module utilizes split operation instead of convolution operation to layer features and concat all Bottleneck outputs, ensuring YOLOv8 lightweight while obtaining richer gradient flow information. In the feature fusion network, YOLOv8 removes the CBS convolution module before the two upsampling operations, and replaces the C3 module used in YOLOv5 with the C2f module. On the prediction end, YOLOv8 abandons the Anchor-based idea and adopts the Anchor-free (Zhu et al.[17]) idea for design. Anchor-free to some extent solves the problem of severe imbalance between positive and negative samples caused by the excessive number of Anchor-based preset anchor boxes and the presence of a large number of negative samples in the background area.

In addition, YOLOv8 also uses Decoupled Head instead of Coupled Head, which decouples regression tasks and classification tasks, to some extent alleviating the conflicts between classification and regression tasks caused by spatial misalignment (Zhang et al.[18]). Although this may reduce the inference speed of the model to some extent, it can improve the detection accuracy of the model. In terms of loss function, YOLOv8 model only calculates classification and regression loss, and no longer calculates confidence loss separately. BCE loss is used for classification loss, and DFL (Distribution Focal Loss) and CIoU are used for regression loss. In YOLOv6, DFL has already been used. Due to the possibility of ambiguity in

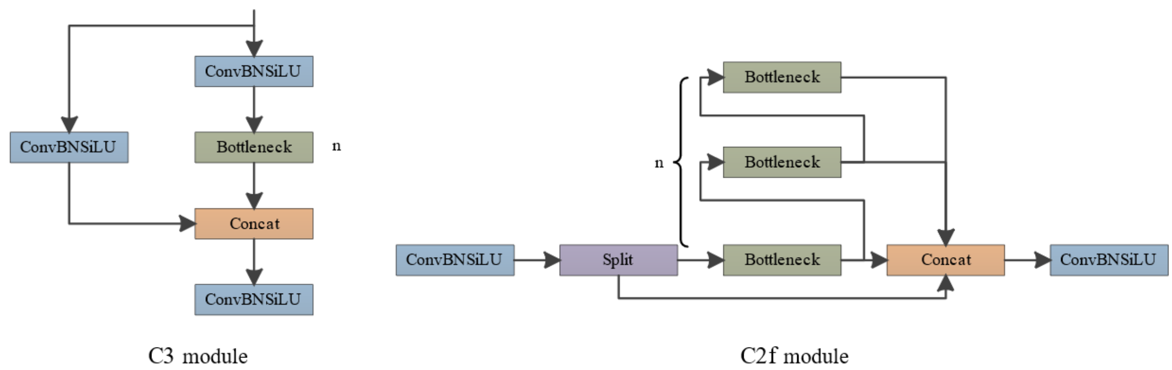

Figure 2 C3 module and C2f module structures

data distribution, DFL simplifies the prediction box positions of continuous distributions to discrete distributions.

3. Improved YOLOv8 Model

3.1. Incorporates CBAM attention mechanism

Although the backbone network of the YOLOv8 model has strong feature extraction capabilities, in natural outdoor environments, aerial images of insulators are susceptible to factors such as lighting, occlusion, complex backgrounds, and the small area of insulator defects, making it difficult to recognize the images. Considering that the spatial features of the input image are significantly weakened after the ninth layer of feature extraction, this paper integrates the CBAM attention mechanism between the fourth C2f module and SPPF module of the YOLOv8 backbone feature extraction network to solve the above problem. The CBAM attention mechanism combines channel attention module (CAM) and spatial attention module (SAM), The feature map maintains the channel dimension unchanged in CAM, compresses the spatial dimension, and makes the model pay more attention to the category information of the image; Keeping the spatial dimension unchanged and compressing the channel dimension in SAM makes the model pay more attention to the position information of the image.

The CBAM attention mechanism strengthens the feature expression of spatial and semantic information in the input feature map, in order to transmit stronger semantic and spatial features during the feature fusion stage, making the YOLOv8 model pay more attention to the characteristics of the insulator itself, while weakening the influence of factors such as lighting, occlusion, and complex background on the detection results. The CBAM attention mechanism structure used in this article is shown in Figure 3.

In CAM, in order to calculate the semantic features of images more efficiently, it is necessary to compress the spatial dimensions of the input image. The CAM structure is shown in Figure 4. SENet (Squeeze and Extraction Network) proposes the use of global average pooling to compress spatial dimensions, and CAM adds the global maximum pooling method on this basis. The specific steps of CAM are as follows: First, the input feature maps are pooled globally to the maximum and globally to the average, so that the size of the feature map changes from W * H * C to 1 * 1 * C. Then, the two feature maps obtained are input into the shared MLP, and the two activated results are obtained through the ReLU activation function. Finally, the two output results of the shared MLP are added and multiplied by the original input feature map through the sigmoid function, so that the image size changes back to W * H * C.

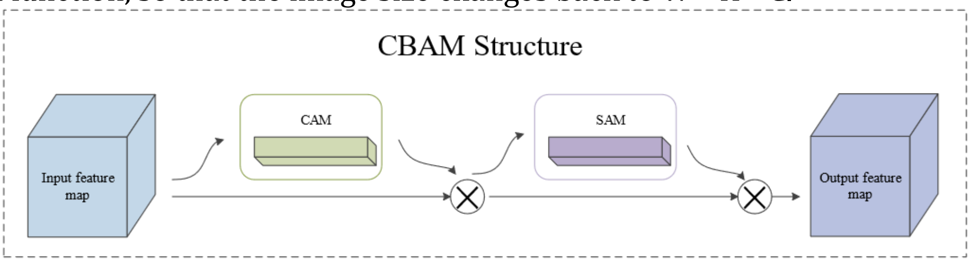

Figure 3 CBAM structure

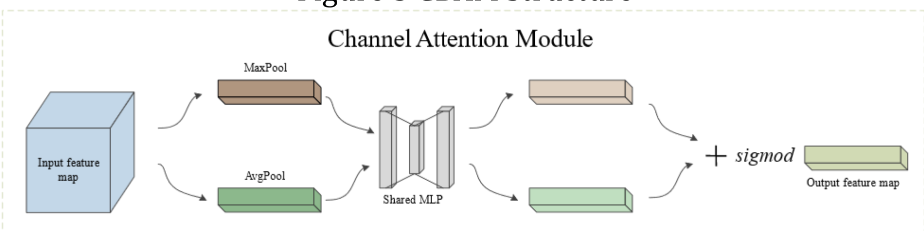

Figure 4 CAM structure

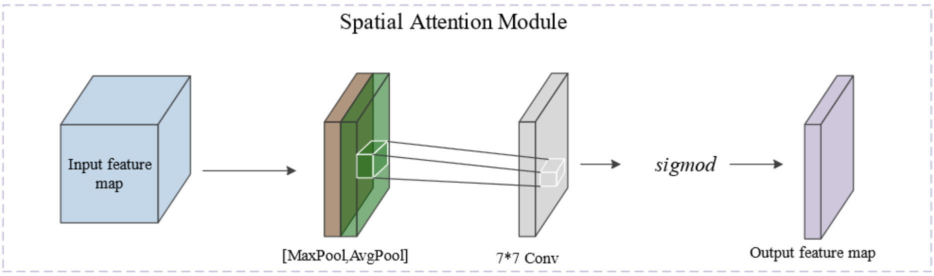

Figure 5 SAM structure

In SAM, in order to calculate the spatial features of images more efficiently, it is necessary to compress the channel dimension of the input image. The output feature map of the channel attention module is used as the input feature map, and the SAM structure is shown in Figure 5. The specific steps are as follows: first, the channel feature map is globally maximally pooled and globally averaged to obtain two W * H * 1 feature maps. These two feature maps are concated, and then dimensionally reduced to a single channel feature map through a convolutional layer with a kernel size of 7 * 7. Finally, the sigmoid function is used to multiply the input feature map, causing the image size to return to W * H * C, and finally, the output feature map of the CBAM attention mechanism is obtained.

In view of the characteristics of complex background and easy occlusion of aerial image of field insulators, CBAM attention mechanism module is fused in YOLOv8 model to enhance the learning of areal feature of image defects such as insulator defect, self explosion, flashover, etc., improve the object detection accuracy, and improve the generalization of the model.

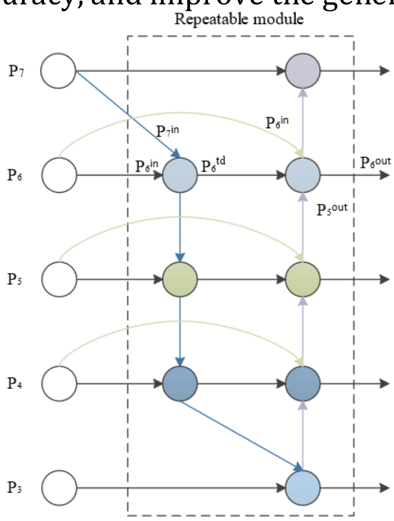

Figure 6 BiFPN structure

3.2. Bidirectional feature pyramid network

Although YOLOv8 replaced all the C3 modules in YOLOv5 with C2f modules on the Neck end, and also removed the convolution operation before two upsampling operations, it still adopts YOLOv5's PANet idea in the feature fusion stage, without fundamentally improving the feature fusion network.

In actual inspection scenarios, due to issues such as shooting angle, shadow occlusion, and insufficient lighting conditions, the image quality of insulator defects may be poor. For such images, the YOLOv8 model cannot extract meaningful features, and the feature fusion effect is poor, which may even affect the learning ability of the model. Therefore, we propose to use the BiFPN structure instead of the original PANet structure to solve the above problems. The BiFPN structure is shown in Figure 6.From Figure 6, it can be seen that the BiFPN structure removes nodes with a single input edge and adds connections between input and output nodes in the same layer. As nodes with only one input edge do not perform feature fusion, their contribution to the feature fusion network is small. Therefore, deleting a single input edge node has little impact on the feature fusion effect. At the same time, deleting a single input edge node can simplify the network structure, reduce computational complexity, and retain the unmerged information of the original nodes. In addition, the BiFPN structure can also perform weighted fusion on the input feature maps. Due to the varying contributions of different input feature maps for feature fusion, simply performing Concat operation is not the best approach. Therefore, we use Fast Normalized Fusion for weighted feature fusion, which allows feature images that contribute significantly to feature fusion to obtain more weights. The weighted calculation method (taking the 6th layer node as an example) is shown in Formula 1.

$$\begin{cases} P_6^{td} = Conv\left(\dfrac{\omega_1 \cdot P_6^{in} + \omega_2 \cdot resize(P_7^{in})}{\omega_1 + \omega_2 + \varepsilon}\right) \\ P_6^{out} = Conv\left(\dfrac{\omega_1^{'} \cdot P_6^{in} + \omega_2^{'} \cdot P_6^{td} + \omega_3^{'} \cdot resize(P_5^{out})}{\omega_1^{'} + \omega_2^{'} + \omega_3^{'} + \varepsilon}\right) \end{cases} \tag{1}$$

Among them, P is the input or output feature of a certain layer node, $Conv$ is a convolution operation, $resize$ is an upsampling or downsampling operation, ω is a weight parameter that distinguishes the importance of different features, ε and is a bias term.

In response to the problem of small object scales in defect areas such as insulator image damage and flashover, this paper proposes a multi-scale feature fusion method using BiFPN structure, which can improve the detection accuracy of YOLOv8 model for small object while reducing network complexity and redundant calculations.

3.3. SIoU loss

YOLOv8 original model used CIoU as the loss function in the regression phase of the prediction box. It calculated the loss of the aspect ratio of the regression box on the basis of DIoU, and comprehensively considered the overlap rate, aspect ratio, and center point distance between the ground truth box and the prediction box (Zheng et al.[19]). This can accelerate the regression speed of the prediction box to a certain extent, making the prediction box more accurate in the regression process. However, in the regression process of the prediction box, when the height and width of the prediction box are proportionally enlarged and shrunk, the regression of the prediction box cannot continue to be optimized, and CIoU also fails to consider that the angle between the prediction box and the ground truth box is also an important factor affecting the regression. Based on the above two considerations, this paper takes SIoU as the loss function in the regression stage of the prediction box. SIoU includes four parts: angle loss, distance loss, shape loss, and intersection merge ratio loss.

In view of a series of problems caused by CIoU as the loss function in the regression phase, using SIoU as the loss function in the regression process of the prediction box can fully consider the angle influence between the prediction box and the ground truth box, and improve the Rate of convergence and regression accuracy of the model, making the whole regression process pay more attention to high-quality anchor boxes. The improved YOLOv8 model structure is shown in Figure 7.

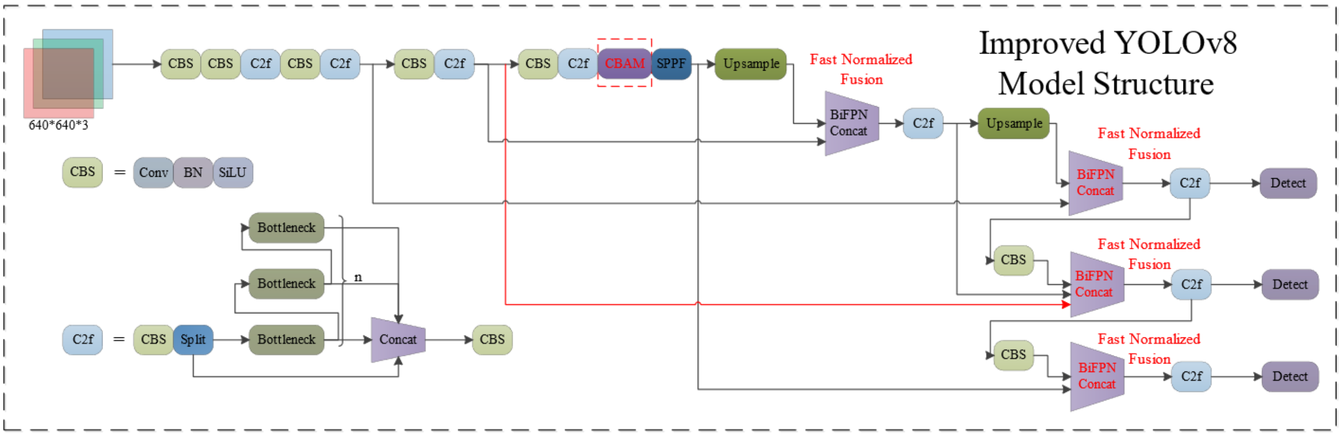

Figure 7 Improved YOLOv8 model structure

4. Experimental results and analysis

4.1. Experimental data processing

The insulator defect samples used in this article are public datasets provided by the Kaggle competition platform. This dataset contains a total of 400 images of insulator damage, self explosion, and flashover. The training set and validation set were randomly divided in an 8:2 ratio. For insulators that have broken or self exploded, label a single insulator, while for insulators that have flashover, only label the flashover area.

4.2. Experimental environment and model configuration parameters

Table 1 Experimental environment configuration

Configuration	Version
OS	Windows10
CPU	Intel i7-9700F
GPU	Geforce 2060 6GB
Python	3.9
CUDA	11.7
Pytorch	1.13

Table 2 Main parameters of the experiment

Parameter Name	Parameter Value
epochs	150
batchsize	16
image size	640*640
optimizer	SGD
lr0	0.01
momentum	0.937
decay	0.0005

In order to verify the effectiveness of the improved YOLOv8 model, Python 1.13 was used as the basic framework for algorithm validation, and CUDA 11.7 was used to accelerate the training process of the model. The specific experimental environment is shown in Table 1.

During the experiment, SGD is used as the optimizer. The initial Learning rate is set to 0.01, the initial momentum is set to 0.937, the weight Attenuation coefficient is set to 0.005, and the IoU threshold is set to 0.5. Other main parameters are shown in Table 2.

4.3. Evaluating indicator

Using precision and recall as basic evaluation indicators, precision is responsible for evaluating the accuracy of model detection, and recall is responsible for evaluating the comprehensiveness of model detection. mAP (mean Average Precision) can be calculated through precision and recall, and its expression is shown in formula 2. Among them, TP is the number of positive samples detected as positive samples, FP is the number of negative samples detected as positive samples, is the number of positive samples not detected, FN is the number of categories, and is the number of categories being tested. Using mAP with an IoU threshold of 0.5 as the final evaluation indicator to evaluate the effectiveness of the improvement on the YOLOv8 model.

$$\begin{cases} P = \dfrac{TP}{TP + FP} \\ R = \dfrac{TP}{TP + FN} \\ AP = \int_0^1 PdR \\ mAP = \dfrac{1}{N}\sum_{i=1}^{N} AP_i \end{cases} \tag{2}$$

4.4. Experimental results

The benchmark model used is YOLOv8, which is compared with the proposed improved YOLOv8 model to verify whether the proposed method can improve the detection accuracy of the model. In terms of detection accuracy, we compared the detection performance of the original YOLOv8 model and the improved YOLOv8 model in the same application scenario and dataset. The specific experimental results are shown in Figure 8.

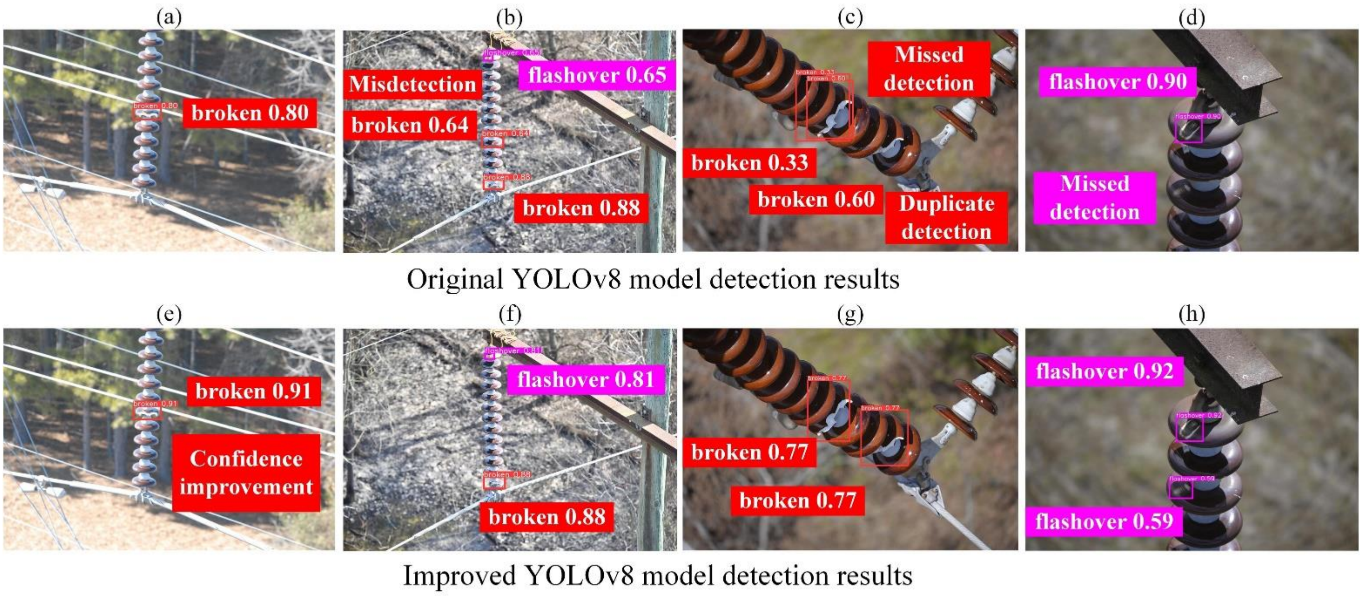

Figure 8 Comparison of original YOLOv8 model and improved YOLOv8 model detection results

The detection results in Figure 8 (a) and Figure 8 (e) show that the improved YOLOv8 model can improve the confidence of object detection, and it can be seen that the confidence in detecting broken insulators has increased by 0.11. The improved YOLOv8 model has improved the feature extraction ability, making object detection more accurate; The detection results in Figure 8 (b) and Figure 8 (f) indicate that the original YOLOv8 model identified intact insulators as broken insulators, resulting in misdetection.

However, the improved YOLOv8 model accurately identified whether insulators had broken faults, effectively avoiding misdetection; The detection results in Figures 8 (c) and 8 (g) indicate that the original YOLOv8 model duplicatedly detects a single object and cannot detect broken insulators with a large occluded area when detecting broken insulators. However, the improved YOLOv8 model can extract the features of the occluded insulator image and accurately identify the insulators with broken faults; The detection results in Figure 8 (d) and Figure 8 (h) indicate that the original YOLOv8 model missed detection when detecting insulation components that were shaded and had flashover faults, while the improved YOLOv8 model weakened the influence of shadows on feature extraction and fusion, effectively improving the accuracy of insulation component fault recognition.

From Figure 8, it can be seen that the improved YOLOv8 model has stronger feature extraction ability, performs better in distinguishing similar features, integrates multi-scale features more thoroughly, has stronger ability to detect small objects, and has lower miss detection and misdetection rates. Even in complex outdoor natural environments, it has good detection performance, good robustness and generalization, and can meet the requirements of unmanned inspection.

In order to objectively verify the effectiveness of the improved YOLOv8 model, this article conducted comparative experiments with mainstream object detection models using the same dataset in the same experimental environment.

Table 3 Comparative experimental results of different object detection models

Model	mAP@50
Faster R-CNN	66.8%
RetinaNet	71.0%
YOLOv5	68.6%
YOLOv6	74.0%
YOLOv8	75.5%
This article	77.8%

Table 3 shows the comparison of detection accuracy between the improved YOLOv8 model and mainstream object detection models such as Faster R-CNN, RetinaNet, YOLOv5, YOLOv6, and YOLOv8. The experimental results show that when the IoU threshold is selected as 0.5, the mAP of the improved YOLOv8 model increases by 11.0%, 6.8%, 9.2%, 3.8%, and 2.3%, respectively, further demonstrating the superior performance of the proposed improved YOLOv8 model.

4.5. Ablation Experiment

Our proposed improved YOLOv8 model improved the original YOLOv8 model on the Backbone end, Neck end, and Head end, respectively, and underwent ablation experiments. The impact of various improvements and their combinations on model performance has been explored, and the results of the ablation experiment are shown in Table 4.

From Table 4, it can be seen that the improved YOLOv8 model has significant improvements in detection accuracy, specifically by 9.4% in Precision and 2.3% in mAP. Through comparison, it can be seen that when CBAM, BiFPN, and SIoU are separately introduced, the detection accuracy of a dataset is improved by 5.4%, 1.2%, and 2.2%, respectively; By combining SIoU with CBAM, CBAM with BiFPN in pairs, the detection accuracy of the same dataset was improved by 4.2% and 3.5%, respectively. Overall, various improvements and their combinations have contributed to the improvement of model detection accuracy, with the CBAM attention mechanism contributing the most to the improvement of model detection accuracy.

Table 4 Results of ablation experiment

Model	Precision	mAP@50
YOLOv8	84.3%	75.5%
YOLOv8+SIoU	86.5%	77.6%
YOLOv8+CBAM	89.7%	76.3%
YOLOv8+BiFPN	85.5%	76.3%
YOLOv8+SIoU+CBAM	88.5%	74.6%
YOLOv8+SIoU+BiFPN	83.3%	77.6%
YOLOv8+CBAM+BiFPN	87.8%	77.5%
This article	93.7%	77.8%

5. Conclusion

This article proposes an improved YOLOv8 insulator defect detection algorithm by analyzing possible problems in actual inspection scenarios, combining attention mechanism with multi-scale feature fusion. Compared with the YOLOv5 model, this algorithm reduces model complexity and shortens detection time; Compared to the original YOLOv8 model, it improves detection accuracy and performs well in overall performance. The specific conclusions are as follows:

1) In response to the complex natural environment in the field and the susceptibility of collected insulator defect images to light and complex backgrounds, the CBAM attention mechanism is integrated into the main feature extraction network to effectively reduce the impact of complex backgrounds on detection accuracy. The unique channel and spatial attention module of the CBAM attention mechanism effectively improve the task of small and medium-sized object detection in insulator fault recognition and object detection in complex backgrounds;

2) In response to the problem of complex and poor feature fusion performance in the original PANet network, a multi-scale feature fusion with BiFPN structure is adopted, fully considering the importance of different feature maps in the feature fusion process. Using weighted fusion method for feature fusion can significantly improve the model's feature fusion ability for objects of different scales, especially small objects;

3) In order to solve the problem that the CIoU cannot continue to optimize the prediction box due to the proportional growth of its length and width in the regression process of the prediction box, SIoU is used instead of CIoU, which can pay more attention to the high-quality anchor box in the regression process, accelerate the Rate of convergence of the model, and further improve the robustness and generalization of the model.

Through experimental verification, it can be seen that the application of the algorithm proposed in this article can timely and accurately identify insulator faults, reduce model complexity, and improve fault identification accuracy. It is of great significance for the detection of insulator faults in transmission and distribution networks and the safe operation of transmission and distribution networks.

References

[1] Liu, Y., Liu, D., Huang, X., Li, C.: Insulator defect detection with deep learning: A survey. IET Gener. Transm. Distrib. 1–18 (2023).

[2]Gouda, O.E., Darwish, M.M.F., Mahmoud, K., Lehtonen, M., Elkhodragy, T.M.: Pollution severity monitoring of high voltage transmission line insulators using wireless device based on leakage current bursts. IEEE Access 10, 53713–53723 (2022).

[3] Wang, Q., Liu, X.: Detection of Self-explosion Defects of Insulators Based on Mask RCNN. Computing Technology and Automation,41(01):52-58 (2022).

[4] Zhao, W., Xu, M., Cheng, X., Zhao, Z.: An insulator in transmission lines recognition and fault detection model based on improved faster RCNN. IEEE Transactions on Instrumentation and Measurement, 70: 1-8 (2021).

[5] Qi, Y., Wu, X,. Zhao, Z., Shi, B., Nie, L.: Bolt defect detection for aerial transmission lines using Faster R-CNN with an embedded dual attention mechanism. Journal of Image and Graphics, 26(11) :2594-2604 (2021).

[6] Yang, S., Deng, D., Zheng, Q.: Ground-glass opacity target detection in CT scans based on improved Faster R-CNN model. Journal of Image and Graphics, 26(09):2171-2180 (2021).

[7] Jia, K., Ma, Z., Zhu, R., Li, Y.: Attention-mechanism-based light single shot multiBox detector modelling improvement for small object detection on the sea surface. Journal of Image and Graphics, 27(04):1161-1175 (2022).

[8] Stefenon, S.F., Singh, G., Souza, B.J., Freire, R.Z., Yow, K.-C.: Optimized hybrid YOLOu-Quasi-ProtoPNet for insulators classification. IET Gener. Transm. Distrib. 17, 3501–3511 (2023).

[9] Cai, Z., Ying, N., Guo, C,. Guo, R., Yang, P.: Research on multiperson pose estimation combined with YOLOv3 pruning model. Journal of Image and Graphics, 26(04):0837-0846 (2021).

[10] Liu, J., Liu, C., Wu, Y., Xu, H., Sun, Z.: An improved method based on deep learning for insulator fault detection in diverse aerial images. Energies, 4(14): 4365 (2021).

[11] Chen, K., Liu, X., Jia, L., Fang, Y., Zhao, C.: Insulator Defect Detection Based on Lightweight Network and Enhanced Multi-scale Feature Fusion. High Voltage Engineering, 1-14[2023-03-27] (2023).

[12] Hao, Y., Liang, W., Yang, L., He, J., Wu, J.: Methods of image recognition of overhead power line insulators and ice types based on deep weakly-supervised and transfer learning. IET Gener. Transm. Distrib. 16, 2140–2153 (2022).

[13] Woo, S., Park, J., Lee, J., Kweon, I.: Cbam: Convolutional block attention module. Proceedings of the European conference on computer vision (ECCV).2018: 3-19 (2018).

[14] Tan, M., Pang, R., Le, Q.: Efficientdet: Scalable and efficient object detection. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.2020: 10781-10790 (2020).

[15] Gevorgyan, Z.: SIoU loss: More powerful learning for bounding box regression. arXiv preprint arXiv:2205.12740 (2022).

[16] Wang, C., Bochkovskiy, A., Liao, H.: YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv preprint arXiv:2207.02696 (2022).

[17] Zhu, C., He, Y., Savvides, M.: Feature selective anchor-free module for single-shot object detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.2019: 840-849 (2019).

[19] Zhang, Z., He, G.: Recognition of Bird Nests on Power Transmission Lines in Aerial Images Based on Improved YOLOv4. Frontiers in Energy Research, 2022: 435 (2022).

[19] Zheng, Z., Wang, P., Ren, D., Liu, W., Ye, R., Hu, Q., Zuo, W.: Enhancing geometric factors in model learning and inference for object detection and instance segmentation. IEEE Transactions on Cybernetics, 52(8): 8574-8586 (2021).