Journal of Electrotechnology, Electrical Engineering and Management (2024) DOI: 10.23977/jeeem.2024.070302
Clausius Scientific Press, Canada ISSN 2560-6697 Vol. 7 Num. 3

Research on Insulator Defect Detection Model Based on Improved YOLOv8

Yang Lu, Xuanrui Hu, Xinzhe Zou, Lei Han

School of Computer and Information Engineering, Heilongjiang University of Science and Technology, Harbin, China

Keywords: YOLOv8 Model, Image Enhancement, CBAM Attention

Abstract: Electricity is the basis for everyday life, and defect detection of insulators of high-voltage transmission lines is the key to ensuring power transmission. In order to overcome the problems of low accuracy of traditional target detection algorithms for small targets, weak representation ability of feature maps and little key information extraction, an improved CBAM attention-based insulator defect detection method CBAM-YOLOv8 based on YOLOv8 was proposed. The core is to apply the combination of Channel Attention Module and Spatial Attention Module to process the channel attention and spatial attention modules respectively for the input feature layers. Experiments show that the AP in the CPLID dataset is improved by 4.7% and the FPS is reduced by 2.7% compared with the original YOLOv8, which proves that the proposed method can maintain a high detection speed while greatly improving the detection accuracy, providing a more effective and safer solution for the detection of high-voltage transmission lines, and greatly reducing the labor cost and operation risk.

1. Introduction

Insulators are a very important part of high-voltage transmission lines, which have a constructive role in support protection and current insulation. Because the insulator is in a very harsh working environment in the field for a long time, it is very easy to produce defects such as self-explosion, and the defects of the insulator will cause safety hazards to the high-voltage transmission line, and may even cause huge economic losses. However, the large scale and complex structure of China's power system, these methods are not accurate, safe, or efficient, and are costly. As a result, grid companies are using deep learning techniques to automatically identify insulator defects [1].

At present, a series of insulator-related object detection algorithms based on meaningful learning already exist. Liao et al [2] integrated the deep residual network Resnet101 into the Faster R-CNN algorithm to detect the defective parts of the insulator, which improved the detection accuracy of the defective parts of the insulator compared with other object detection algorithms, but the computational cost was heavy, resulting in a low detection speed. Yao and Qin [3] used the improved YOLOv3 algorithm to detect insulator defects, and replaced the original IoU loss function with a new loss function GIoU, which improved the accuracy of insulator defect detection and had a better detection speed compared with the original YOLOv3 algorithm without increasing the computational cost, but still did not solve the problem of missed detection. Wang et al [4] proposed an insulator detection algorithm with a good regional suggestion network, which improved the detection accuracy of the network by further improving the ResNeSt network and combining with the multi-scale regional suggestion network RPN, but the detection speed was still low.

Based on the in-depth analysis of the existing insulator defect detection methods, this paper contributes a fast YOLOv8 detection algorithm fused with CBAM module. By adjusting the network structure and training strategy, the algorithm effectively improves the detection speed and accuracy, and controls the computational cost. Experiments on the CPLID dataset show that the proposed algorithm is superior to the existing technology, and provides an efficient scheme for the automatic detection of insulator defects in high-voltage transmission lines, which is helpful for power grid security and reduces economic losses.

2. The principles of improving the YOLOv8 model

2.1 The network structure of improved YOLOv8

YOLOv8 is an object detection model that includes the input layer, backbone network, network neck, and network header. It divides the image into grids, each of which predicts multiple bounding boxes and object probabilities. The model processes the images through the convolutional neural network to generate object detection results. Since insulators may appear in different states due to aging, contamination or damage, and complex backgrounds may interfere with detection, YOLOv8 is improved by introducing CBAM attention mechanism, so that the model can focus more on the key parts of the image, enhance the detection ability of defective insulators, and improve the overall performance. The network structure is shown in Figure 1.

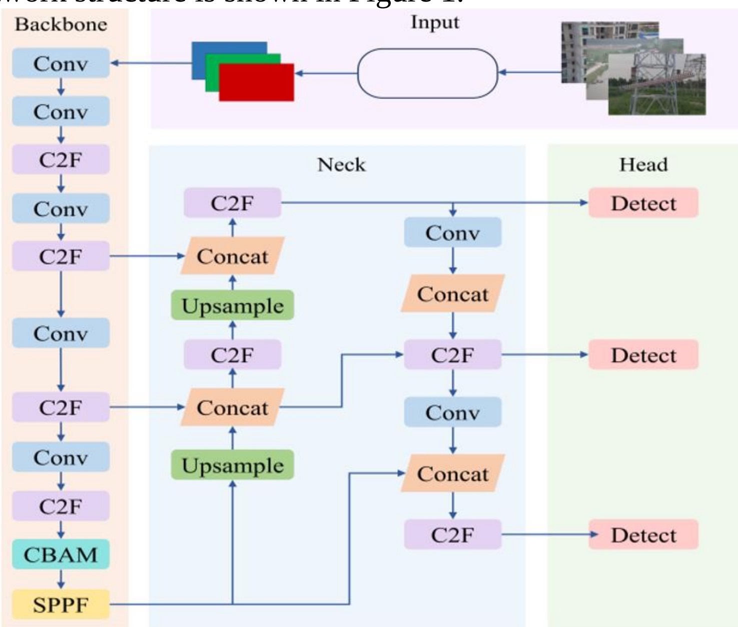

Figure 1: Improved YOLOv8 network structure

2.2 The principles of CBAM attention

CBAM combines the attention mechanism of two dimensions: feature channel and feature space. The core lies in the combination of the Channel Attention Module and the Spatial Attention Module, and the channel attention and spatial attention modules are processed respectively for the input feature layers. Through its staged attention mechanism, the feature channels are weighted first, and then the spatial dimensions are benighted, so as to achieve a deep understanding and representation of complex features. Its staged attention mechanism can make the model pay more attention to the key features of the insulator, and effectively improve the model's ability to identify the unique geometric shape and structural details of the insulator. The network structure is shown in Figure 2.

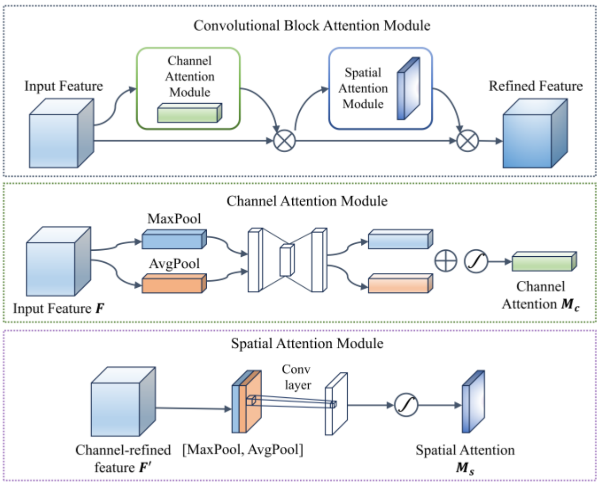

Figure 2: CBAM attention structure

The Cam Attention Mechanism consists of two main parts: Chanel Attentin and Spatiar Attentin. Here is the formula for calculating these two attention modules in Cam:

$$F_{avg} = AvgPool(F) \tag{1}$$

$$F_{max} = MaxPool(F) \tag{2}$$

where F is imported feature maps, F_{avg} indicates average pooling feature plot, $\mathrm{F_{max}}$ represents feature plot after maximum pooling

$$\mathrm{MLP}\left(\mathrm{F_{avg}}\right) = \mathrm{W_1} \cdot \mathrm{RELU}\left(\mathrm{W_0} \cdot \mathrm{F_{avg}}\right) \tag{3}$$

$$\mathrm{MLP}(\mathrm{F_{max}}) = \mathrm{W_1} \cdot \mathrm{RELU}(\mathrm{W_0} \cdot \mathrm{F_{max}}) \tag{4}$$

where $\mathrm{W_0}$ and $\mathrm{W_1}$ is weight matrix in MLP, RELU indicates activate the function

$$\mathrm{M_c(F)} = \sigma(\mathrm{W_1} \cdot \mathrm{ReLU}(\mathrm{W_0} \cdot AvgPool(F)) + \mathrm{W_1} \cdot \mathrm{ReLU}(\mathrm{W_0} \cdot MaxPool(F)))$$

where σ is : Sigmoid function to map the combined attention weights between 0 and 1

$$F'_{avg} = AvgPool(F') \tag{5}$$

$$F'_{max} = MaxPool(F') \tag{6}$$

where F' is : Feature map weighted by channel attention

$$F'_{cat} = Concat\left(F'_{avg}, F'_{max}\right) \tag{7}$$

$$F'_{conv} = Conv(F'_{cat}) \tag{8}$$

where Conv is : Convolution operation, using an appropriately sized convolutional kernel (7x7)

$$\mathrm{M_s}(F') = \sigma\left(\mathrm{Conv}^{7\times7}([AvgPool(F'); MaxPool(F')])\right) \tag{9}$$

where $\mathrm{M_c}(F')$ is : Final spatial attention weight map

$$F' = \mathrm{M_c(F)} \odot \mathrm{F} \tag{10}$$

where $\odot$ is : Element-by-element multiplication

$$F'' = \mathrm{M_s}(F') \odot F' \tag{11}$$

where F'' is : The final output feature map contains the weighting of channel and spatial attention mechanisms

3. Results

3.1 Model training environment and evaluation indicators

The experiment is based on the 64-bit operating system Ubuntu 20.04 LTS, and the deep learning framework is PyTorch v1.10.0, and experimental environment configuration is shown in Table 1.

Table 1: Experimental Environment

Experimental environment	Version model
Operating System	Ubuntu 20.04 LTS
CPU	Intel(R) Core(TM) i7-10700K
CPU	NVIDIA GeForce RTX3060
Deep Learning Framework	PyTorch v1.10.0
IDE	Pycharm v2023.2.4

Each indicator of specific definitions are shown in (12) to (13).

$$\mathrm{R} = \frac{TP}{TP+FN} \tag{12}$$

where R is recall, TP indicates the number of positive samples that were correctly predicted, FN represents the number of positive samples that were incorrectly predicted as negative

$$\mathrm{P} = \frac{TP}{TP+FP} \tag{13}$$

where P is precision, FP indicates the number of negative samples that were incorrectly predicted as positive

$$\mathrm{AP} = \int_0^1 P(R)dR \tag{14}$$

where AP is : average accuracy, which is a measure of the model's performance

$$\mathrm{mAP} = \frac{\sum AP}{N} \tag{15}$$

where mAP is when there are multiple detection targets, the mean average accuracy is selected, N indicates the total category of the detected target number

$$\mathrm{FPS} = \frac{1}{t} \tag{16}$$

where FPS is the frame rate per second, which is used to evaluate the detection speed of the model, t indicates the time it takes to process an image

3.2 CPLID DataSet

The China Power Line Insulator Dataset (CPLID) is an open-source image dataset designed for power line insulator defect detection. This dataset consists of normal insulator images captured by drones and synthetic defective insulator images, aiming to promote the application of deep learning in power facility maintenance, due to the insufficient number of datasets, the dataset is expanded to 13659 images using image processing techniques such as flipping, contrast adjustment and symmetry transformation, including 11498 in the training set and 3600 in the test set. Figure 3 shows some examples of insulator images in the CPLID dataset.

Figure 3: Examples of the CPLID dataset images

3.3 Analysis of experimental results

The images in the CPLID dataset face certain challenges due to their small defective targets, complex and changeable backgrounds, and many interfering targets. In this paper, we propose to compare CBAM-YOLOv8 with object detection algorithms such as Faster R-CNN [5]. Faster R-CNN [5] uses the regional recommendation network for object detection, which shows high detection efficiency. The SSD [6] algorithm, which acts as a single-stage detector capable of object detection on feature maps at different scales, was also included in the comparison. RetinaNet [7] is another important benchmark for comparison by introducing focus loss to solve the category imbalance problem. The DetectoRS [8] algorithm is also considered, which improves detection performance through recursive feature pyramids and switchable dilated convolution. In addition, the original YOLOv8 algorithm will also be involved in the comparison. Table 2 shows the comparison of the detection accuracy and speed of these algorithms on the CPLID dataset.

Table 2: The algorithm in this paper is compared with other algorithms on the CPLID dataset

Method	AP	FPS
Faster R-CNN	81.6%	4.3
SSD	84.9%	18.6
RetinaNet	87.4%	16.5
DetectoRS	91.6%	2.9
YOLOv8	89.8%	30.5
CBAM-YOLOv8	94.5%	27.8

As can be seen from the table 2, after a series of improvements, the AP of the proposed method CBAM-YOLOv8 achieves the best results on the CPLID dataset, reaching 94.5%, which is 4.7% higher than the original YOLOv8 algorithm. However, due to the increase of network structure complexity and the increase of computational cost while improving the network, the FPS is 2.7% lower than that of the original YOLOv8 algorithm but the decrease is small, so compared with other object detection algorithms, the proposed method still has a high detection speed and can well meet the requirements of the insulator defect detection task, and the proposed method has obvious advantages in the detection accuracy, which proves the accuracy of the method in insulator image defect detection.

In this article, the CPLID dataset is processed and divided into training set: test set: validation set=8:1:1, and the loss value and accuracy of the training set and the validation set are visualized. The visualization of CBAM-YOLOv8 algorithm is shown in Figure 4.

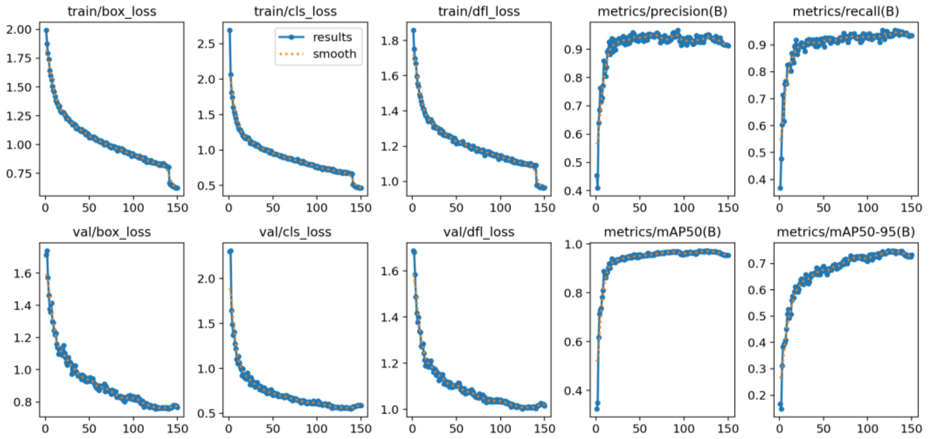

Figure 4: Visualize the results

In all the presented graphs, it is observed that the loss values consistently diminish as the number of iteration cycles increases. This indicates that the model's performance in terms of error rate is improving over time. Concurrently, the accuracy values exhibit an initial phase of ascent, which suggests that the model is progressively becoming more proficient at correctly classifying or predicting the outcomes. Following this initial rise, the accuracy values reach a plateau, signifying that the model has stabilized and is no longer experiencing significant gains in accuracy, despite the ongoing reduction in loss. This stabilization typically occurs once the model has approached an optimal solution where further improvements are minimal or incremental.

Accuracy and recall are two key indicators to measure the performance of a classification model, where accuracy reflects how accurate the model is in predicting the positive class, that is, how many of the samples predicted as positive are true positive classes, while recall measures the model's ability to identify positive classes, that is, how many of the samples that are actually positive are correctly predicted by the model, and these two indicators can evaluate the good or bad of a model. The performance indicators are shown in Figure 5.

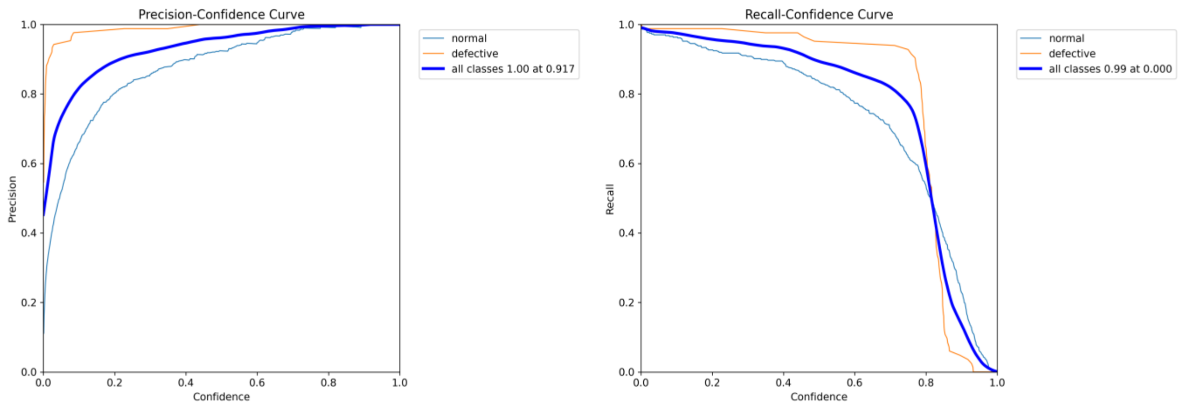

Figure 5: Performance indicators

The performance index of a model in the training process shows a significant downward trend through the CBAM-YOLOv8 algorithm, which indicates that the model continues to learn and optimize during the training process, effectively reducing the prediction error. At the same time, the accuracy or precision curve showed a steady upward trend and reached a high level, which indicates that the model performs well in recognition or classification tasks and has high reliability. The recall rate is also maintained at a high level, indicating that the model performs well in capturing relevant instances, and the CBAM-YOLOv8 algorithm shows efficient learning ability and excellent performance.

4. Conclusions

The defect detection of insulator images by using object detection algorithm has been the concern of many scientific research teams and researchers. In this paper, the traditional object detection algorithm focuses on the low detection accuracy, weak feature characterization ability and key information extracted by the traditional object detection algorithm in this paper, an insulator defect detection method based on CBAM attention mechanism was proposed, CBAM-YOLOv8.

The combination of Channel Attention Module and Spatial Attention Module is applied to process the input feature layer with channel attention and spatial attention module respectively, so that the model can be in the background eliminate interference in complex images, focus on key information, and avoid false detections; Finally, the experimental results show that the CBAM-YOLOv8 proposed in this paper has obvious advantages, which fully proves the effectiveness of the proposed method in the defect detection task of insulator images. In the next research work, we will try to study more lightweight networks to meet the requirements of detection accuracy and speed, and make the model more flexible and convenient.

References

[1] Zhu Yongkun, Wang Junliang, et al. Analysis of flashover fault causes and preventive measures for composite insulators [J]. Inner Mongolia Electric Power Technology, 2021, 39(6): 82-86.

[2] Liao G P, Yang G J, Tong W T, et al. Study on power line insulator defect detection via improved faster region-based convolutional neural network[C]. Proceedings of the 2019 IEEE 7th International Conference on Computer Science and Network Technology. Dalian, China: IEEE, 2019: 262-266.

[3] Yao L N, Qin Y Y. Insulator detection based on GIOU-YOLOv3[C]. Proceedings of the 2020 Chinese Automation Congress. Shanghai, China: IEEE, 2020: 5066-5071.

[4] Wang S Q, Liu Y F, Qing Y H, et al. Detection of insulator defects with improved ResNeSt and region proposal network [J]. IEEE Access, 2020, 8: 184841-184850.

[5] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[C]. Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press, 2019: 91-99.

[6] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multiBox detector[C]. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016: 21-37.

[7] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]. Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 2999-3007.

[8] Qiao S Y, Chen L C, Yuille A. DetectoRS: Detecting objects with recursive feature pyramid and switchable atrous convolution[C]. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, TN, USA: IEEE, 2021: 10208-10219.