Improved YOLOv4 Power Insulator Fault Detection

Zhen Zhang¹, Shuaihua Kong¹* and Kun Peng¹
¹School of Electrical Engineering, Zhengzhou University, Zhengzhou, Henan, 450001, China
*Corresponding author’s e-mail: ksh1985357098@gs.zzu.edu.cn

Abstract. In order to achieve accurate real-time monitoring of power insulators on transmission lines, an improved YOLOv4 network model is proposed to detect power insulator faults. Experimental results show that the mAP value of the improved YOLOv4 network model reaches 93.01%, and the detection frame rate reaches 44 frames/s. Compared with the YOLOv4 and Faster R-CNN models, it can be concluded that the proposed model can better meet the needs of power insulator fault detection for transmission lines.

1. Introduction
As an important power equipment component of power transmission lines, insulators play a vital role in the operation of power systems[1]. Insulators are exposed to harsh outdoor environments for a long time and are also fault-prone components. Relevant data shows that insulator component failures occur most frequently in power grid equipment. Therefore, the research on insulator detection is of great significance[2].

In order to achieve the goal of real-time, high-precision and high-efficiency detection of fault images of transmission line insulators, this paper selects the YOLOv4 target detection model that can achieve real-time detection to detect transmission line insulators and improves the model[3]. The analysis of experimental results shows that after the improvement, the model improves the fault detection performance of insulators on transmission lines.

2. YOLOv4 network structure
The feature extraction network of the YOLOv4 algorithm is different from that of YOLOv3. It uses a new feature extraction network CSPDarkNet53 instead of Darknet53 in the YOLOv3 model. In the CSPDarkNet53 feature extraction network structure, the activation function of DarknetConv2D uses the Mish activation function to replace the original leaky-ReLU activation function, and the convolution block in the network structure is changed from DarknetConv2D_BN_Leaky to DarknetConv2D_BN_Mish. The YOLOv4 network structure introduces the SPP (Spatial Pyramid Pooling) structure and the PANet (Path Aggregation Network) structure in the feature pyramid. The SPP module uses four different scales of maximum pooling($13 \times 13$, $9 \times 9$, $5 \times 5$, $1 \times 1$) for processing, and converts input images of different sizes into fixed-size output, which can increase the perception field of view, separate contextual features, and improve scale invariance. The PANet network structure uses a bottom-up sampling strategy to complete image feature extraction, and realizes feature aggregation with FPN (Feature Pyramid Networks)[4], which can improve the detection effect of small target objects. The YOLOv4 network structure diagram is shown in Figure 1.
3. Improved YOLOv4 algorithm

3.1. Bounding box regression loss function selection

In the actual target detection, the bounding box regression loss function is the evaluation index of the
detection effect of the target detection algorithm. At present, the YOLO series of algorithms often use
Intersection-over-Union (Intersection-over-Union) as the bounding box regression loss function, and its
calculation formula is shown in formula (1).

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}$$  \hspace{1cm} (1)

In the above formula: A represents the true frame, and B represents the predicted frame.

The intersection ratio IoU represents the degree of overlap between A and B. During the training
data set process, it is found that there will be disjoint situations between A and B, resulting in failure to
train. In order to solve the above-mentioned problems, Hamid Rezatofighi and Nathan Tsoi et al.
proposed the GIoU loss function in 2019[5]. GIoU not only pays attention to the intersection of A and
B, but also other disjoint parts, which helps to reflect the degree of overlap between A and B. The
calculation formula is shown in formula (2).

$$\text{GIoU} = \text{IoU} - \frac{|C \setminus A \cup B|}{|C|}$$  \hspace{1cm} (2)

In the above formula: C represents the minimum frame area of A and B.

Although the loss function GIoU has a certain degree of improvement in detection accuracy
compared to IoU, there is a situation where the two boxes overlap indefinitely, that is, GIoU = IoU.
Therefore, Z Zheng and P Wang et al. proposed DIoU and CIoU bounding box regression loss functions
in 2020[6]. The DIoU loss function can directly optimize the distance between the two boxes. For the
situation where the two boxes overlap infinitely, the DIoU convergence speed Faster.

The CIoU loss function takes into account the aspect ratio of the predicted frame to fit the aspect
ratio of the target frame on the basis of DIoU, making the target frame regression more stable. The
calculation formula of the CIoU loss function is shown in formulas (3), (4) and (5).

$$\text{CIoU} = \text{IoU} - \frac{\alpha^2}{c^2} - \alpha v$$  \hspace{1cm} (3)

$$\alpha = \frac{v}{(1 - \text{IoU}) + v}$$  \hspace{1cm} (4)

$$v = 4 \left( \frac{\arctan \frac{w_{gr}}{h_{gr}} - \arctan \frac{w}{h}}{\pi^2} \right)^2$$  \hspace{1cm} (5)
In the above formula: \( \rho \) represents the Euclidean distance between the prediction frame and the center point of the real frame; \( c \) represents the diagonal distance of the smallest closure area including the prediction frame and the real frame at the same time; \( b \), \( w \) and \( h \) represent the center point, width and height of the prediction frame respectively; \( g_{tb} \), \( g_{tw} \) and \( g_{th} \) respectively represent the center point, width and height of the real frame.

The CIoU loss function function is to measure the degree of overlap between the prediction box and the real box. In actual optimization, CIoU cannot be directly used as the loss value (loss) operation. If CIoU is directly optimized, it will lead to the difference between the prediction box and the real box. The degree of overlap is getting lower and lower. Therefore, CIoU needs to be processed, and the Loss calculation formula is shown in formula (6).

\[
Loss = 1 - IoU + \frac{\rho^2 (b, g_{tb})}{c^2} + \alpha v
\]  

(6)

3.2. Introduce attention mechanism

In the actual transmission line insulator detection, the target of the insulator to be detected is often accompanied by a complicated background, which makes it difficult to identify the target's fault characteristics, and it is easy to miss the detection. Therefore, the introduction of attention mechanism into insulator fault detection can not only focus on the important features of insulators, but also suppress or ignore other irrelevant feature information in the background, which can effectively improve the accuracy of insulator fault detection. Sanghyun Woo and Jongchan Park et al. proposed CBAM (Convolutional Block Attention Modul) in 2018[7], which is the attention mechanism module. This module has two sequential sub-modules, namely the channel attention module and the spatial attention module. In this paper, the CBAM module is added to the YOLOv4 network model, which will further improve the detection effect of the network model. The schematic diagram of the CBAM structure is shown in Figure 2.

3.3. Network structure improvement design

The YOLOv4 network model uses CSPDarknet53 as the backbone feature extraction network. CSPDarknet53 extracts multiple feature layers for target detection, and there are 3 a priori boxes for each feature layer, among which 3 prior boxes correspond to large, medium and small target objects respectively. The CSPDarknet53 structure takes the input insulator image through a series of convolution operations to obtain a total of 5 convolutional layers, layer1, layer2, layer3, layer4, and layer5. The CBAM module is connected to the 3 effective convolutional layers in the CSPDarknet53 feature extraction network, as shown in Figure 3. Won the positions marked with ①, ②, and ③. The YOLOv4 network uses the PAN network structure and the FPN network structure to extract features from three effective convolutional layers. Among them, the PAN network structure is sampled from top to down for feature extraction, and the FPN network structure is sampled from bottom to up for feature extraction. Feature fusion realizes the full extraction of features in the three feature layers. The three sizes after fusion are 19×19, 38×38, 76×76 and correspond to YOLO-1, YOLO-2, YOLO-3 in the figure, and finally, through the three YOLO layers, the prediction results such as the position, category, and confidence of the insulator candidate frame are obtained to achieve End-to-end detection. The improved network structure diagram is shown in Figure 3.
4. Experiment and result analysis

4.1. Experiment procedure

This paper uses the PASCAL VOC 2007 data format and uses the labelImg labeling tool to label the insulator data set. During the labeling process, as many insulators in the image (label insulator) and insulator fault locations (label faulty) are labeled as much as possible. The experiment was completed on the PC side, and the computer GPU was NVIDIA GeForce RTX 2080 Ti. The basic parameter settings of the model during the training process are 100 epochs, the learning rate is 0.0014, the batch size is 64, and the number of iterations is 100. In the actual training process, the parameters are constantly adjusted to achieve the optimization of the model. The Loss curve is shown in Figure 4. After 70 iterations, the curve gradually stabilizes.

![Loss curve](image)

The use of sections to divide the text of the paper is optional and left as a decision for the author. Where the author wishes to divide the paper into sections the formatting shown in table 2 should be used.
4.2. Loss function comparison experiment

In order to analyze and compare the influence of different loss functions on the algorithm model, this paper uses four different loss functions of IoU, GIoU, DIOU and CIoU to verify on the YOLOv4 algorithm. Use the self-made insulator data set (image size) to analyze the training result after training for 100 epochs. The evaluation index of the training result is expressed by the value of mAP. The comparison results of different loss functions are shown in Table 2. The mAP change curve of the training process is shown in Figure 5.

Table 1. Experimental comparison table of different loss functions

| Algorithm name | Loss function | mAP/%  |
|----------------|---------------|--------|
| YOLOv4         | IoU           | 86.88  |
|                | GIoU          | 88.31  |
|                | DIOU          | 89.87  |
|                | CIoU          | 91.82  |

Figure 5. mAP change curve of different loss functions

It can be seen from Table 2 that different loss functions have certain differences in the detection performance of the YOLOv4 network model. Among them, the detection ability of the IoU loss function is poor, and the detection ability of the CIoU loss function is better and the detection accuracy reaches 91.82%. And from the change curve of loss function mAP in Fig. 5, CIoU converges faster than the other three loss functions and is easier to reach stability. Therefore, choosing CIoU loss function as the loss function of the model in this paper can effectively improve the detection performance of faulty insulators.

4.3. Comparison experiment of different algorithms

This paper compares the improved algorithm with the mainstream SSD, YOLOv4, Faster R-CNN three target detection algorithms on the self-made insulator data set for insulator fault detection comparison experiments. The detection effect of the experiment takes mAP and FPS as the evaluation index of the algorithm, and the comparison of the detection effect of different algorithms is shown in Table 2.

Table 2. Comparison of detection effects of different algorithms

| Algorithm name            | mAP/% | FPS/(f/s) |
|---------------------------|-------|-----------|
| SSD                       | 76.35 | 79        |
| YOLOv4                    | 80.62 | 53        |
| Faster R-CNN              | 85.54 | 6         |
| The algorithm of this article | 93.01 | 44        |
From the experimental results in Table 4, it can be seen that the detection accuracy of the algorithm in this paper has reached 93.01%, which is not only 7.47% higher than the detection accuracy of the Faster R-CNN algorithm, but the detection frame rate is still 7.33 times that of the algorithm. Although the detection speed of the algorithm in this paper is lower than that of SSD algorithm and YOLOv4 algorithm, its detection accuracy is higher than SSD and YOLOv4 algorithm, and the detection frame rate is greater than 25 frames/s to meet the real-time requirements of insulator video monitoring. In summary, the target detection algorithm in this paper can better fulfill the requirements of insulator fault detection in terms of detection accuracy and detection speed. The detection effect of this algorithm is shown in Figure 6.

5. Conclusion
Aiming at the problems of low detection accuracy and slow detection speed when directly using the YOLOv4 target detection algorithm to detect power insulator images, this paper proposes an improved YOLOv4 network insulator fault detection model. Experiments show that compared with the YOLOv4 and Faster R-CNN algorithms, the mAP value of this algorithm has increased by 12.39% and 7.47%, respectively, and the FPS has reached 44 frames/s. It can be seen that the improved YOLOv4 algorithm can accurately detect whether the insulator is faulty under the premise of ensuring real-time performance.

Acknowledgments
This work was supported by Zhengzhou High-tech Zone Double Innovation Base n (20180397A).

References
[1] Lv Z N. A survey of common faults analysis and detection methods for transmission line[J]. Automation and Instrumentation, 2020, No.243(01):167-170+174.
[2] Zhang H K, Li J Y, Zhang B. Foreign Object Detection on Insulators Based on Improved YOLOv3[J]. Electric Power, 2020, 053(002):49-55.
[3] Bochkovskiy A, Wang C Y, Liao H. YOLOv4: Optimal Speed and Accuracy of Object Detection [J/OL]. 2020, https://arxiv.org/abs/2004.10934.
[4] Liu S, Qi L, Qin H, et al. Path Aggregation Network for Instance Segmentation[C]. computer vision and pattern recognition, 2018: 8759-8768.
[5] Rezatofighi H, Tsoi N, Gwak J, et al. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression[C]. computer vision and pattern recognition, 2019: 658-666.
[6] Zheng Z, Wang P, Liu W, et al. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression[C]. national conference on artificial intelligence, 2020.
[7] Woo S, Park J, Lee J, et al. CBAM: Convolutional Block Attention Module[C]. european conference on computer vision, 2018: 3-19.