ORIGINAL PAPER

Research on the algorithm of detecting insulators in high-voltage transmission lines using UAV images

Yachao Si[1] · Jiajie Gao[1] · Mingzhan Zhao[1] · Xiaojun Xu[1]

Abstract

Insulator defect detection is an important task of power line inspection, but because the insulator defect target occupies a small area in the transmission line image, and the background is complex and changeable, the accuracy of insulator defect detection is low. Based on previous research results and existing problems, an insulator defect detection algorithm based on YOLOv5 is proposed. Based on the original YOLOv5, C2f module is proposed to replace C3 module in backbone part of YOLOv5, Wise-IoU loss function is introduced, and SimAM attention mechanism is added to head part of the original YOLOv5. As for the above improvement, the method can accurately and effectively identify the insulator and accurately judge whether there is a defect problem under the complex background. We conducted training and comparison experiments on different models on the CPLID insulator public data set and the custom data set of Roboflow and designed ablation experiments to analyze the effects of each improvement part. The experimental results show that the detection accuracy of the improved YOLOv5n model in the insulator defect target detection task reaches 97.7% and mAP50-95 reaches 78.7%. Compared with the original YOLOv5n model, the accuracy of the improved YOLOv5 model is increased by 0.6% and mAP50-95 by 1.1 percentage points. The improved algorithm can be better applied to the target detection of insulator defects in UAV. Considering the high probability that mobile devices do not support high-performance Gpus and the simple network structure of YOLOv5 is more lightweight than the YOLOv8 model, we chose the 5 version of YOLO instead of the latest YOLOv8.

Keywords Deep learning · YOLOv5 · Wise-IoU · SimAM

1 Introduction

As a basic component of transmission lines, insulators play an important role in electrical isolation and mechanical support. However, because they are often in a harsh natural environment, insulators are vulnerable to external factors and cause damage, which may lead to power line interruption in serious cases. Therefore, insulator detection has become an important part of transmission line inspection that cannot be ignored. In the process of insulator defect detection, there will be some complicated situations, the common situation is that the insulator is damaged (as shown in Fig. 1a), and there may also be occluder on the insulator, which may be animals (including birds, reptiles, insects, etc., Fig. 1b shows the snake on the insulator). The presence of animals on the

insulator is generally regarded as normal. If the cover is a bird's nest, it is considered a defect that needs to be addressed (Fig. 1c shows this).

Insulator defect detection is a hot research field, and many experts at home and abroad are exploring effective detection methods. With the development of deep learning and the progress of computers, insulator defect detection methods based on deep learning have gradually become the mainstream. As an application example of object detection, the algorithm of insulator detection can follow the mature algorithm of object detection. These methods use deep neural networks to extract the features of insulator images, and classify and locate defects. Literature [1] uses the improved faster RCNN algorithm to detect insulator defects on power lines, but only one category of experimental results is not convincing. In literature [2], based on generative adversarial network (GAN), U-Net network is used to detect insulator defects, and the detection accuracy is significantly increased, but the background is complex and changeable, and data enhancement is lacking. Literature [3] uses the three fea-

✉ Xiaojun Xu
  xxj_31@163.com

[1] Hebei University of Architecture, Qiaodong District, Zhangjiakou 075000, China

Fig. 1 Insulator condition diagram

(a)

(b)

(c)

(d)
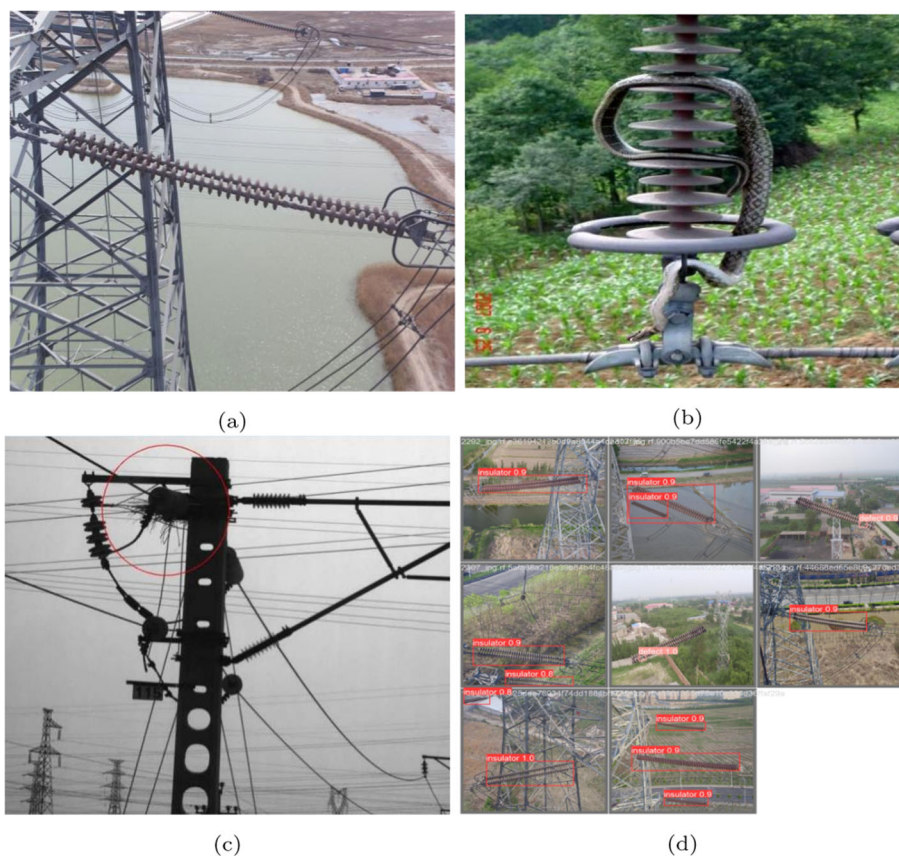
tures of contrast, variance and mean value for comprehensive analysis to diagnose whether the insulator has self-explosion. However, the current method lacks adaptability in the area division. In literature [4], the detected insulator region was processed into blocks, and then the distance measurement between chromaticity was used for statistical analysis. The method is sensitive to environmental changes and has insufficient sensitivity. Literature [5–9] is the experimental effect achieved by the latest research on the improvement of YOLO model. While the single performance is improved, the accuracy and mAP obviously have relatively large errors. In the literature [10], insulator defect detection based on deep learning convolutional autoencoder (DCAE) was used to encode and decode the insulator image, and then whether there was a defect was determined according to the reconstruction error and anomaly score. This method did not need a large amount of labeled data and could effectively extract image features. When the accuracy is improved, the image is processed by slicing, and the appropriate threshold is set to distinguish the normal and abnormal images, which makes the early workload quite large, and there are too many influencing factors such as image quality and noise, resulting in poor detection results. Literature [11] uses a deep learning insulator defect detection method based on the framework of two-stage target detection Fast-RCNN +FPN. Compared with the one-stage

network, the two-stage network structure can better deal with the problem of multi-scale and small targets. At the same time, the improvement of deep learning algorithm leads to more calculation and poor real-time performance. The global RoI method leads to the large size of the feature map, occupies more memory and computing resources, and reduces efficiency. Therefore, the development of more efficient and adaptive target detection technology has become an urgent task for the challenges and needs unique to the diversity and complex background of insulators.

2 Related work

In this study, the YOLOv5 model was used for insulator defect detection. The YOLOv5 model was improved in view of problems such as high missed or false detection rate, diversity of insulators, and inability to accurately identify small defective insulators in insulator defect image detection. The improvement work is as follows: (1) C2f of YOLOv8 is used to replace C3 of backbone part of YOLOv5. The structure of C2f module is inspired by the design of C3 module and ELAN module, which can realize the bidirectional fusion of features and thus realize the lightweight of features. (2) Adding the SimAM [12] attention mechanism to the head

part of YOLOv5 can effectively enhance the expression ability of features and improve the adaptability of multi-scales by adjusting the attention of feature maps before each detection head. (3) The Wise-IoU [13] loss function is introduced to dynamically update the normalization factor, which can ensure that the gradient gain remains at a high level as a whole, so as to optimize the training effect of the model.

YOLOv5 and YOLOv8 are both advanced target detection algorithms in the field of deep learning. They both adopt the backbone-neck-head network architecture, but there are some differences in details. YOLOv5 has fewer model parameters and a smaller amount of computation, making it more suitable for operating in resource-constrained environments, such as mobile devices or edge computing platforms. On both the CPU and GPU, YOLOv5 is able to achieve higher frames per second (FPS), which is an important advantage for applications that require real-time processing. YOLOv5 does not need to design a large number of anchor boxes, and it simplifies the detection process by directly predicting the center point of the target, which is especially advantageous in small target detection. For specific detection tasks, such as insulator detection, these characteristics make YOLOv5 a more suitable choice. Although YOLOv8 may have some improvements in some aspects, YOLOv5 still has significant advantages in lightweight, speed, stability, and adaptability to specific scenarios.

YOLOv5, as a relatively advanced network architecture in the YOLO series and a single-stage object detection network, has achieved good results in detection accuracy and speed. The YOLOv5 algorithm model mainly consists of three parts: the backbone network, the feature fusion module (Neck), and the detection module (Detect). This structure provides an efficient and accurate solution for image object detection, and its detailed structure is shown in Fig. 2.

(1) A concise backbone network: mainly used for extracting image features, using CSPNet [14] to segment feature maps and cross connect them at different stages, thereby reducing computational complexity and improving the learning ability of feature maps; it includes the focus module, which reduces the width and height of the input image by half through slicing operations, while increasing the number of channels by four times. This allows for more effective extraction of image detail features without increasing computational complexity; introduced adaptive anchor box calculation, which can automatically adjust the size and proportion of anchor boxes based on the training dataset.

(2) Feature Fusion Module (Neck): It is used to fuse feature maps of different scales, YOLOv5 adopts the PANet (path aggregation network) structure, which enhances information flow and feature utilization by adding paths between low-level and high-level feature maps; in the neck section of YOLOv5, there is not only top-down feature fusion, but also a bottom-up path, which can better preserve high-resolution detail information while utilizing deep semantic information; it also includes a spatial pyramid pooling (SPP) module, which can pool features at different scales, making the model more robust to the size of the input image.

(3) Detect: For locating the position of the target and predicting the class of the target, YOLOv5 uses anchor boxes to predict the bounding box of the target. In each anchor box, YOLOv5 predicts the class and position offset of the target. During the detection process, there will be many overlapping prediction boxes. YOLOv5 uses the NMS algorithm to remove these overlapping prediction boxes and only keeps the prediction boxes that are most likely to contain the target.

3 Methodology

This article takes YOLOv5 as the baseline model and proposes YOLOv8's C2f to replace the C3 in the backbone part of YOLOv5. SimAM attention mechanism is added to the head part of the original YOLOv5, and the Wise-IoU loss function is introduced. The improved model structure diagram is shown in Fig. 3.

3.1 C2f module concept and introduction effect

The C2f module is a feature fusion layer based on the C3 module, which draws on the ideas of ELAN and aims to improve gradient flow information and lightweight capabilities. The C2f module is mainly divided into two key parts: SPP (spatial pyramid pooling) and PAN (path aggregation network). Through the synergistic effect of these two parts, the C2f module can introduce more effective feature representations into the neural network. Firstly, SPP achieves effective feature extraction for targets of different sizes by constructing a pyramid structure with pooling layers of different scales. This can enable the network to have better perception ability and recognize objects of different sizes; next is PAN, which mainly solves the problem of information fusion between feature maps of different scales. The PAN module adopts multiple cross-stage partial network fusion (CSP) modules to fuse information from shallow and deep feature maps. This can increase the perception range of the network and improve the accuracy of object detection.

During the process of insulator defect detection using drones, due to the irregular size of defects and the varying distances between drones and targets, there may be issues with accuracy not meeting expectations in actual insulator defect detection. Therefore, this improvement is based on the orig-
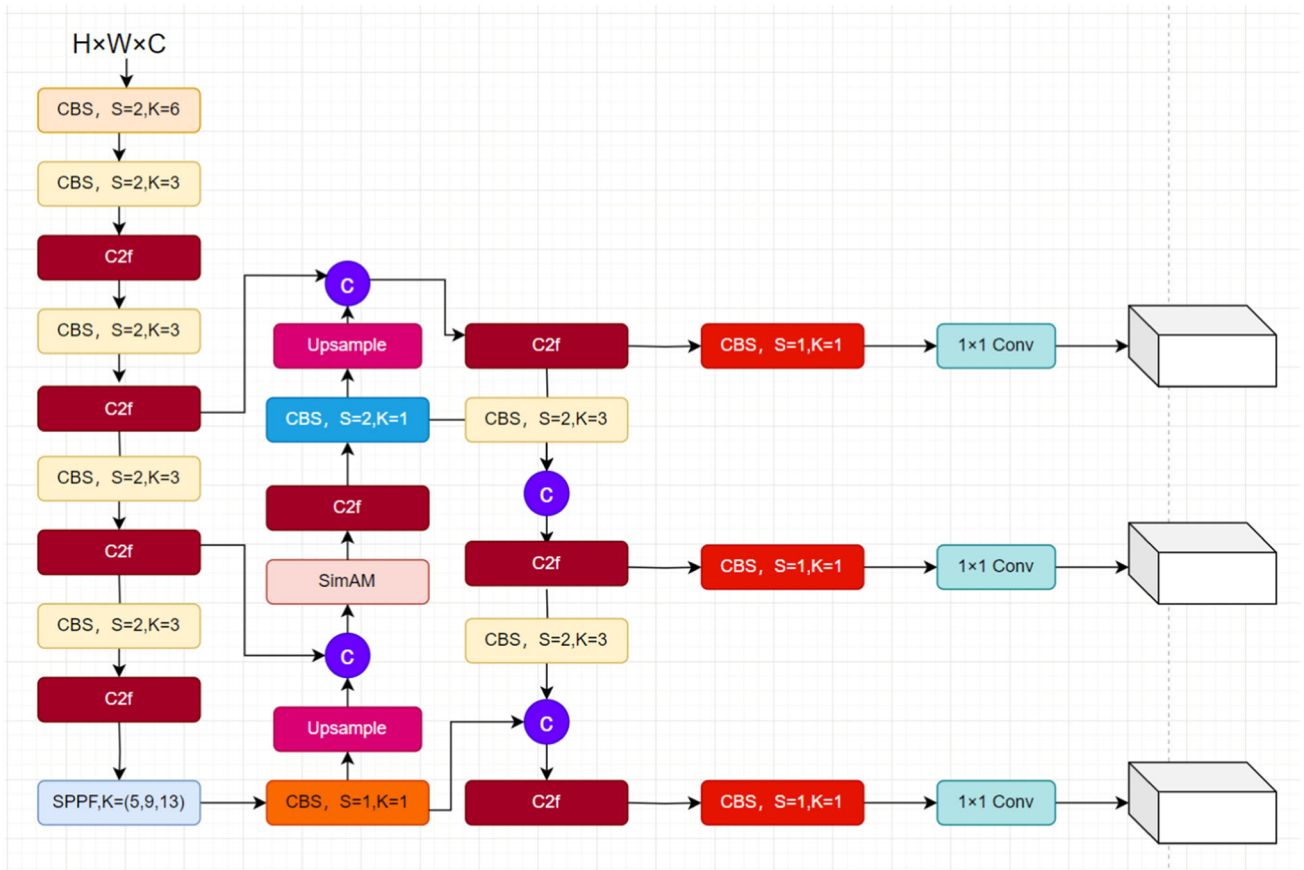
Fig. 2 YOLOv5 algorithm structure diagram

inal C3 of YOLOv5, replacing the original C3 module of YOLOv 5 with C2f. By adopting the C2f module, the model can leverage its efficient feature merging and channel partitioning mechanism to reduce computational requirements and accelerate inference speed. In addition, the design of the C2f module has also optimized the extraction and integration of multi-scale features, thereby improving the performance of the model in identifying small targets. The C2f structure diagram is shown in Fig. 4:

3.2 Adding wise-IoU loss function

In the field of insulator defect detection, the loss function plays a crucial role. The main goal is to improve the accuracy of the detection model by optimizing the position error between the predicted frame and the enclosing frame of the actual insulator defect. By reducing the difference between the predicted frame and the real frame, the optimization of the loss function enables the model to more accurately identify and locate defects on the insulator. Because the training data inevitably contains low-quality examples, geometric measures such as distance and aspect ratio will aggravate the penalty for low-quality examples and reduce the generalization performance of the model. A good loss function should

reduce the penalty of geometric measures when the anchor frame and the target frame coincide well, but more intervention training will make the model have better generalization ability. Therefore, the Wise-IoU v1 version is adopted as a new loss function in this study to balance the model training results pushed by different masses and obtain more accurate detection results. The parameters of Wise-IoU are shown in Fig. 5:

IoU loss is defined as:

$$L_{\text{Iou}} = \frac{\left(\text{B} \cap B^{\text{gt}} \right)}{\left(\text{B} \cup B^{\text{gt}} \right)} \tag{1}$$

In the formula: B is the predicted bounding box, and B^{gt} is the real bounding boxes.

The distance loss formula for the Wise-IoU loss function is:

$$R_{\text{Wise-IoU}} = \exp \left[\frac{\left(x - x_{\text{gt}} \right)^2 + \left(y - y_{\text{gt}} \right)^2}{\left(W_{\text{g}}^2 + H_{\text{g}}^2 \right)^*} \right] \tag{2}$$

$$L_{\text{Wise-IoU v1}} = R_{\text{Wise-IoU}} \, L_{\text{Iou}} \tag{3}$$

Fig. 3 YOLOv5 algorithm improvement structure diagram
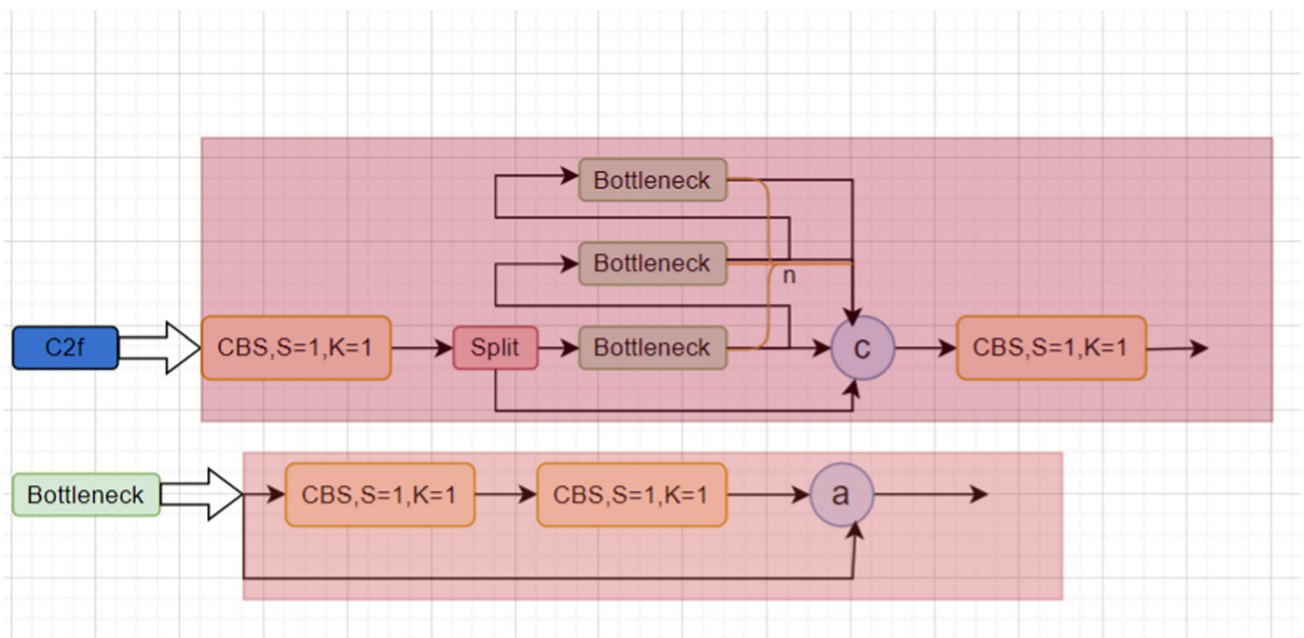

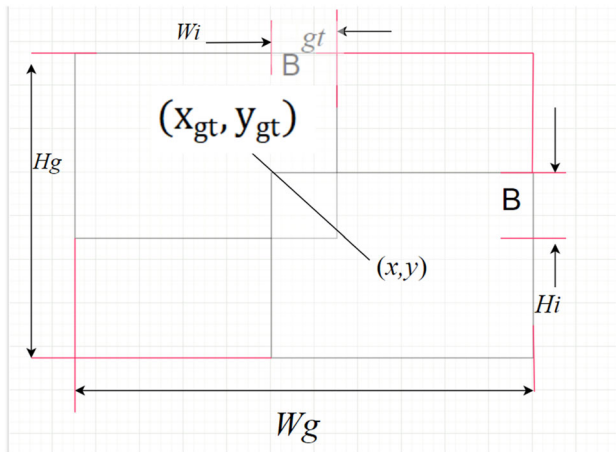
Fig. 4 C2f structure

Fig. 5 Wise-IoU parameters diagram

In the formula: As shown in Fig. 5, W_g and H_g belong to the minimum enclosed box size; x, y are the coordinate of the center point of the prediction box; x_{gt} and y_{gt} are the coordinates of the center point of the real box.

$R_{\text{Wise-IoU}} \in [1, e)$, this will significantly enlarge the L_{Iou} of the ordinary quality anchor box. $L_{\text{Iou}} \in [0, 1]$, this will significantly improve the $R_{\text{Wise-IoU}}$ of high-quality anchor frames, and significantly reduce its focus on the distance to the center point when the anchor box and target box overlap well.

To prevent $R_{\text{Wise-IoU}}$ from generating gradients that hinder convergence, separate W_g and H_g from the computational graph (superscript * indicates this operation) because it effectively eliminates factors that hinder convergence.

After the above improvements, the YOLOv5n model has improved its recognition ability and robustness during the image preprocessing stage, further enhancing its feature extraction and fusion capabilities.

3.3 Adding SimAM attention mechanism

The attention mechanism is a novel attention module discovered based on the combined processing of human visual and brain signals. For the task of detecting insulator defects with similar background environments, research has found that introducing attention mechanisms in the head can effectively remove irrelevant interference. Therefore, this article introduces a parameter-free attention module SimAM. This attention mechanism is implemented by an energy function, which is related to spatial inhibition in visual neuroscience. In other words, in visual processing, neurons that exhibit significant spatial inhibitory effects should be given higher priority (i.e., importance). The simplest way to find these neurons is to measure the linear separability between a target neuron and other neurons. Based on these neuroscience findings, we defined the following energy function for each neuron:

$$e_t(w_t, b_t, y, x_i) = \frac{1}{M-1} \sum_{i=1}^{M-1} (-1 - (w_t x_i + b_t))^2 + (1 - (w_t t + b_t))^2 + \lambda w_t^2 \quad (4)$$

The weight and deviation obtained by solving formula (4) are shown in formula (5) and formula (6):

$$W_t = -\frac{2(t - u_t)}{(t - u_t)^2 + 2\sigma_t^2 + 2\lambda} \quad (5)$$

$$b_t = \frac{1}{2}(t + u_t) w_t \quad (6)$$

Among them, u_t and σ_t represent mean and variance, and the SimAM attention mechanism structure diagram is shown in Fig. 6:

The formula of the energy function indicates that the lower the energy, the greater the difference between the neuron and the surrounding neurons, thus increasing its importance. Based on this concept, the output formula of SimAM (similarity attention module) can be derived from this. Among them, E represents the sum of the minimum energy in all spatial and channel dimensions, and A is the input feature. After applying the sigmoid function to the outliers that suppress attention weights, multiplying them with the corresponding elements of the input feature matrix yields the final output:

$$\widehat{A} = \text{Sigmod}\left(\frac{1}{E}\right) \cdot A \quad (7)$$

This experiment places SimAM in the head section to improve the model's attention to important features in the image, thereby enhancing the performance of object detection. SimAM adaptively adjusts the information in the feature map to enable the model to focus more on useful features for the current insulator defect detection task.

4 Experiments and results

4.1 Experimental environment construction

The experimental parts involved in this article were all completed under the agreed experimental environment, which consists of hardware and software. The specific parameters are shown in Tables 1 and 2.

4.2 Dataset collection and image enhancement

The experiment used the Chinese Power Line Insulator Dataset (CPLID) public dataset, with a total of 848 images and an input image size of 640 × 640. Due to the limited variety and quantity of insulator defect datasets, in order to improve the accuracy of model training results and prevent
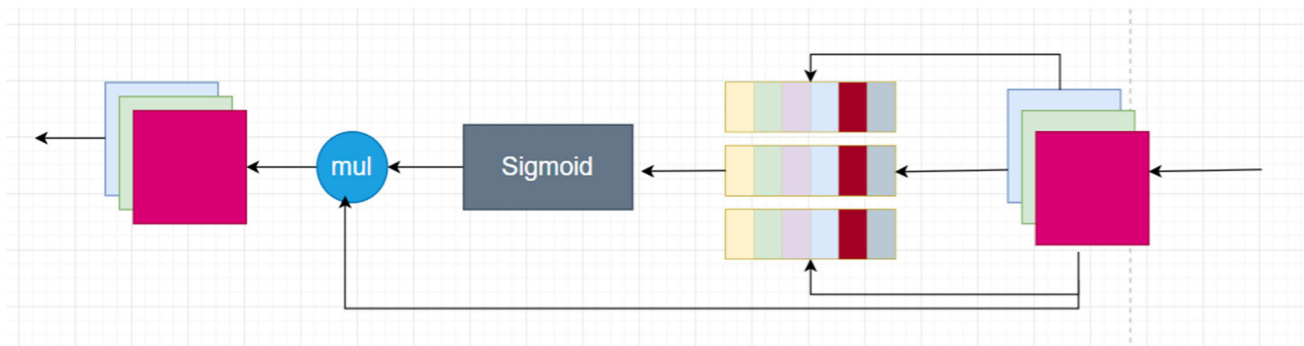
Fig. 6 SimAM attention mechanism structure

Table 1 Configuration of experimental hardware environment

Experimental environment	Configuration
System	Windows 10
CPU	Intel(R) Xeon(R) Silver 4210 CPU @ 2.20GHz
	2.19GHz
RAM	256 GB
GPU	GeForce GTX 3090
Memory	24 GB

Table 2 Experimental software environment configuration

Experimental environment	Configuration
CUDA	12.1
Python	3.9
Anaconda	2.0.3
Pytorch	1.8.1

overfitting during the process, this study enhanced the dataset with images and performed the following enhancement operations:

Flip: Flip the insulator image clockwise and counterclockwise by 45 degrees each, for a total of 90 degrees.

Brightness: Use Roboflow to enhance and decrease the brightness of the image by 30 degrees each.

Saturation: Increase or decrease the saturation of the insulator dataset image by 30 degrees, totaling 60 degrees.

After the above data augmentation operations, Roboflow is used to convert the data set format into YOLO format output, and the data set is automatically re-labeled with insulator and defect locations. The final dataset was expanded to 5581 images in total. Before the model training, the data set was divided into 4906 images in the training set, 506 images in the test set, and 169 images in the validation set.

In order to carry out comparative experiments to prove that all the innovations of this improvement are feasible, a custom data set is used at the same time. The experiment uses the custom data set ins defect on Roboflow. It contains 2089 images collected from different angles, illumination, scale, lighting, background, and insulator types, and contains 2089 images of real environment occlusion and scale change. When training the model, the dataset was divided into 1497 images as the training set and 395 images as the validation set.

4.3 Comparative analysis of experimental results

4.3.1 C2f contrast experiment and effect analysis

A comparative experiment was conducted on YOLOv5n model to verify the performance of the improved algorithm. There were four sets of test results, including accuracy, recall rate, mAP and mAP50-95 before and after the improvement.

The first group is the initial YOLOv5n model; Group 2 refers to the change of C3 module to C2f module on the basis of group 1, and group 1 and group 2 are comparative experiments conducted on public data sets. Group 3 and Group 4 are experimental comparisons on custom data sets on Roboflow. The P–R curve of the comparison experiment is shown in Fig. 7. Table 1 compares the performance indicators. Table 3 shows the results of C2f comparison experiment.

As can be seen from Table 3 of the comparative experiments, after C2f module is used to replace C3 module, the recall rate of the improved model on the public data set increases by 0.4%, mAP by 0.3%, and mAP50-95 by 2.5% and improved accuracy on custom datasets by 0.9% and mAP50-95 by 1.5%. The comparative experimental results show that the performance of replacing C3 with C2f on the original YOLOv5n model is improved, thus verifying the feasibility of the improvement and the effectiveness of insulator defect detection.

4.3.2 Loss function comparison experiment and experimental results

In order to comprehensively evaluate the performance of the improved model in this study, comparative experiments were designed with different loss functions on a common dataset.
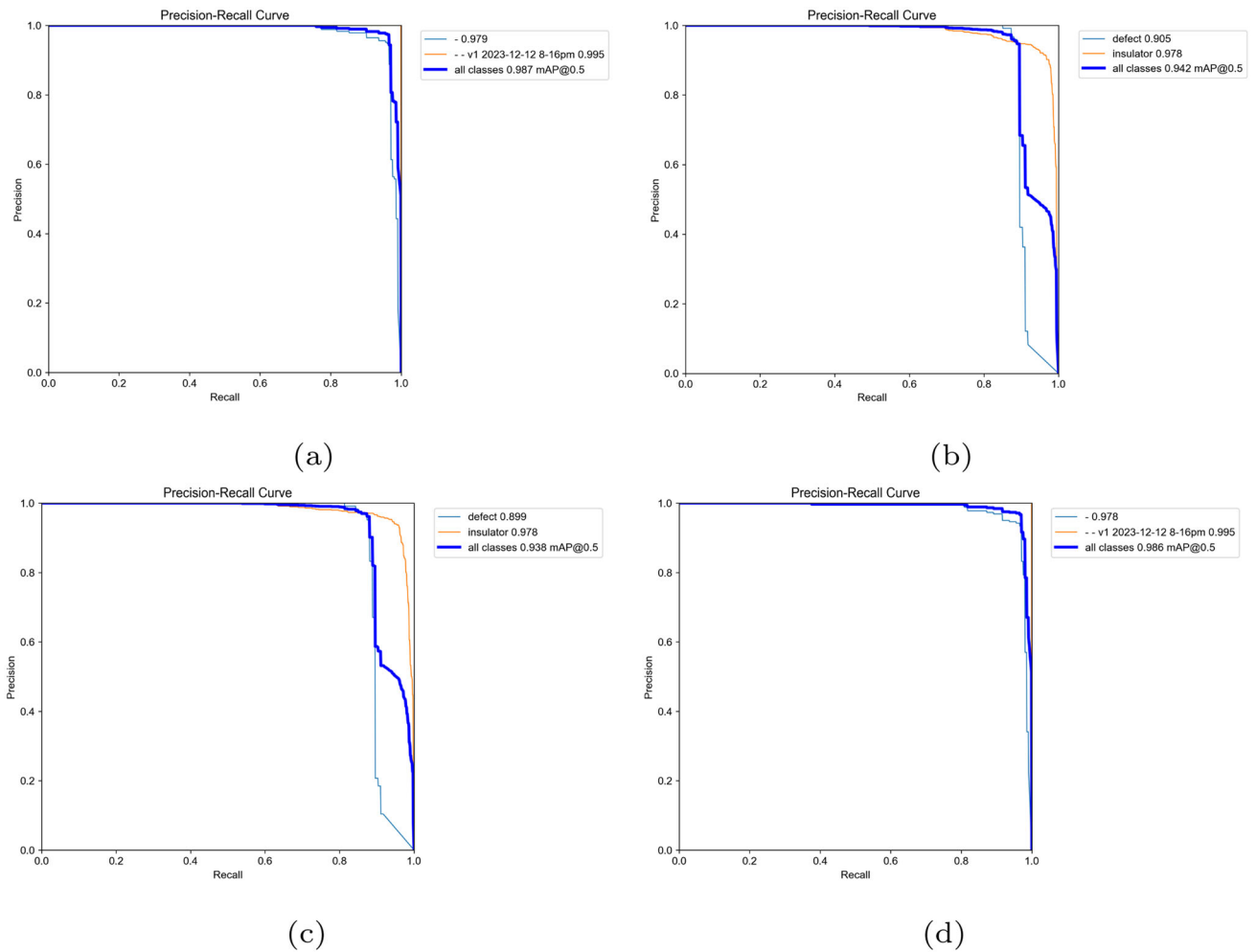
(a)

(b)

(c)

(d)

Fig. 7 Precision–Recall curve

Table 3 Results of C2f comparison experiment

Model	Accuracy	Recall rate	mAP50	mAP50-95
1. YOLOv5n	97.1	98.0	98.7	77.6
2. YOLOv5n+C2f	96.4	98.4	99.0	80.1
3. YOLOv5n	94.9	91.1	94.2	68.6
4. YOLOv5n+C2f	95.8	90.6	93.8	70.1

Table 4 Comparative experimental results of loss function

Model	Accuracy	Recall rate	mAP50	mAP50-95
1. YOLOv5n	97.1	98.0	98.7	77.6
2. YOLOv5n+Giou	96.6	98.3	98.6	77.2
3. YOLOv5n+Siou	96.9	97.5	98.4	77.7
4. YOLOv5n+Diou	96.6	97.6	98.5	78.6
5. YOLOv5n+Eiou	96.5	98.3	98.9	78.6
6. YOLOv5n+Wise-loU	96.7	97.8	98.9	78.2

To investigate the impact of the Wise-IoU loss function on algorithm performance in the YOLOv5n model, this paper embeds five common loss functions into the algorithm network YOLOv5n. YOLOv5n defaults to using the Ciou loss function, forming a comparative experiment between the five loss functions proposed in Table 4 and the model before and after improvement.

For the detection of insulator defect images, different loss functions were added to the YOLOv5n model under the same hardware environment. According to Table 4, after introduc-

ing the Wise-IoU loss function, the mAP50 of the model reached a maximum of 98.9, which is currently the best performance compared to the other loss functions introduced this time. This verifies the feasibility of selecting Wise-IoU as the loss function. The results of the YOLOv5n model with the introduction of the Wise-IoU loss function are shown in Fig. 8.
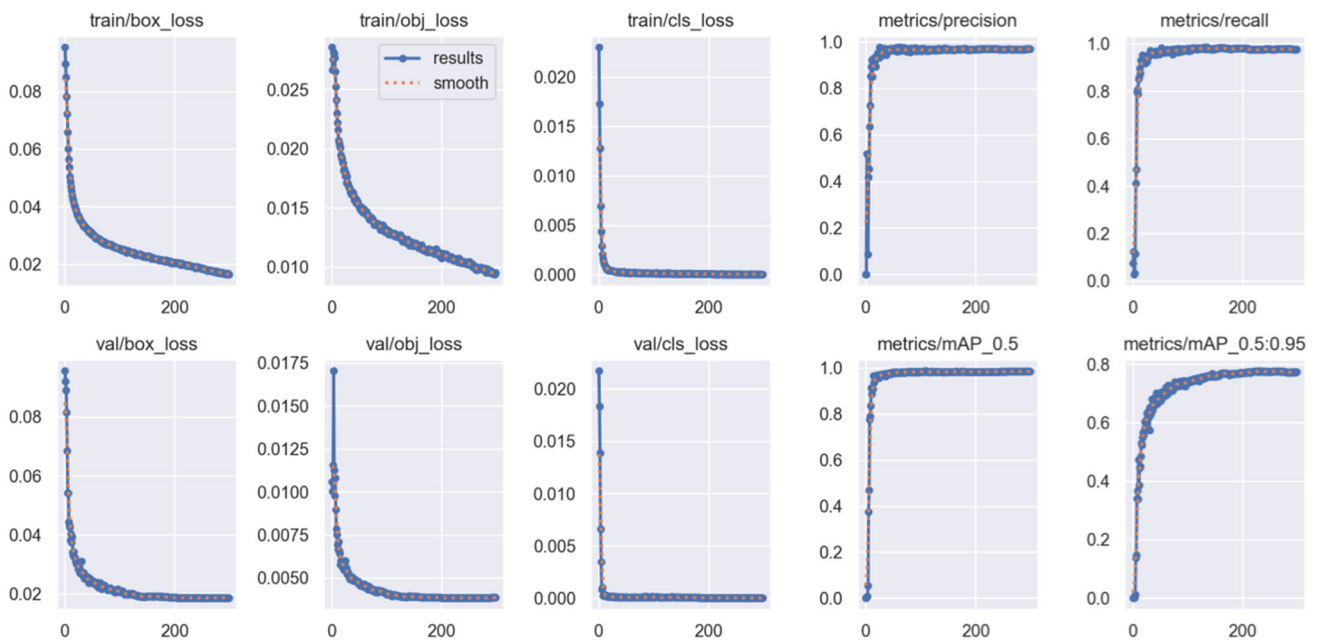
Fig. 8 Improved model results

Table 5 Experimental results of attention mechanism comparison

Model	Accuracy	Recall rate	mAP50	mAP50-95
1.YOLOv5n	97.1	98.0	98.7	77.6
2.YOLOv5n+ SEAttention	97.7	97.0	98.9	77.9
3.YOLOv5n+MHSA	96.6	97.8	98.6	77.9
4.YOLOv5n+GAM	96.8	98.0	98.9	78.5
5.YOLOv5n+ TripleAttention	97.5	98.0	98.9	78.4
6.YOLOv5n+ ShuffleAttention	97.1	97.7	98.4	77.8
7.YOLOv5n+SGE	97.5	98.0	98.6	77.2
8.YOLOv5n+SK	97.1	98.4	98.8	77.4
9.YOLOv5n+ SimAM	96.3	98.0	98.7	78.2

4.3.3 Comparative experiment and result analysis of attention mechanism

At present, there are thirteen commonly used attention mechanisms in the improved model. This study selected seven comparative experiments with SimAM attention mechanisms to verify the effectiveness of this improvement. The comparative experiment is shown in Table 5.

After comparing the final algorithm model with different attention mechanisms such as GAM, SK, and SGE embedded in the public dataset provided in this article, it can be concluded that although the SimAM attention mechanism does not have the highest performance among several attention comparisons, its performance in recall, mAP50, and

mAP50-95 is relatively balanced and above average, with stable performance. The other types of attention perform best in a certain aspect, but the overall performance difference is too large, which is inconsistent with the original intention of this improvement concept. Therefore, it is still best to introduce the SimAM attention mechanism for performance.

4.3.4 Ablation experiment

In order to verify the effectiveness of replacing C3 module of YOLOv 5 with C2f module, adding SimAM attention to backbone part, and Wise-IoU loss function improved in this paper, ablation experiments are carried out to evaluate the influence of each improved module on the detection algorithm in this paper. The ablation experiment uses the original YOLOv5n experimental results as the benchmark, and the performance of the experiment on the public data set is shown in Table 6, and the performance under the custom is shown in Table 7.

Table 6 Ablation results 1

Improving network structure	Accuracy	Recall rate	mAP50	mAP50-95
YOLOv5n	97.1	98.0	98.7	77.6
YOLOv5n+C2f	96.4	98.4	99.0	80.1
YOLOv5n+SimAM	96.3	98.0	98.7	78.2
YOLOv5n+Wise-IoU	96.7	97.8	98.9	78.2
YOLOv5n+C2f+ SimAM+Wise-IoU	97.7	97.5	98.7	78.7

Table 7 Ablation results 2

Improving network structure	Accuracy	Recall rate	mAP50	mAP50-95
YOLOv5n	94.9	91.1	94.2	68.6
YOLOv5n+C2f	95.8	90.6	93.8	70.1
YOLOv5n+SimAM	94.9	91.1	94.2	68.6
YOLOv5n+Wise-IoU	95.4	91.1	94.3	69.4
YOLOv5n+C2f+ SimAM+Wise-IoU	95.9	90.4	93.9	71.0

It can be seen from Tables 3 and 7 that the original YOLOv5n can obtain 94.9% precision, 91.1% recall rate, $94.2\% mAP50 - 95$ and $68.6\% mAP50 - 95$ results on the ins defect data set. After gradually adding the three improved modules, the detection indicators are basically improved. It shows that each module contributes to the task of insulator defect detection on transmission lines and also validates the rationality of taking enhancing the gradient information flow of the model and reducing the influence of noise features, reducing the competitiveness of high-quality anchors, and reducing the harmful gradients generated by low-quality samples as the starting point. In the public data set and the custom data set, YOLOv5n model has significantly improved the accuracy and mAP50-95 of the models after the intro-duction of the three innovations. Thus, the feasibility and efficiency of the improved model can be illustrated, and insulator defects can be detected with higher accuracy. The detection results are shown in Fig. 9.

Firstly, the accuracy of YOLOv 5n is improved from 97.1% to 97.7%, and the recall rate is improved from 98.0% to 98.4% under the public data set after replacing the C3 module of YOLOv5N with C2f. The mAP50 increased from 98.7% to 99%, and the MAP50-95 value increased by 0.2%, which proves that the improved model helps to improve the object detection accuracy of the model in low light environment. Secondly, the mAP50-95 value is increased by 0.6% after introducing the SimAM attention mechanism, which proves that the model with the SimAM attention mechanism can strengthen the model's attention to the target feature by simplifying the calculation process without significantly increasing the computational burden. After introducing the Wise-IoU loss function, the mAP is increased by 0.2% and the mAP50-95 is increased by 0.6%, which proves that the improved model can optimize the regression of the bounding box by dynamic focusing mechanism, so as to improve the detection accuracy. Finally, when the model integrates three improvements, the accuracy is significantly improved to 97.7%, and the mAP50-95 is improved to 78.7%. The above data clearly verify the feasibility of this improvement for the insulator defect detection task.


Fig. 9 Insulator defect detection results

Table 8 Comparison of results of different detection algorithms

Detection algorithm	Accuracy	Recall rate	mAP
FasterNet-YOLOv5 [5]	97.57		93.3
YOLOv4+SE-CSPDarkNet+loU-aware [6]	94.31	90.31	93.76
YOLOv5s+ODConv +CAAttention[7]	95.3	90.3	95.0
YOLOv5-GSEM [8]	91.1	82.4	87.8
EfficientNet-YOLOv5s [9]	92.5	97.9	98.5
Ours	97.7	97.5	98.7
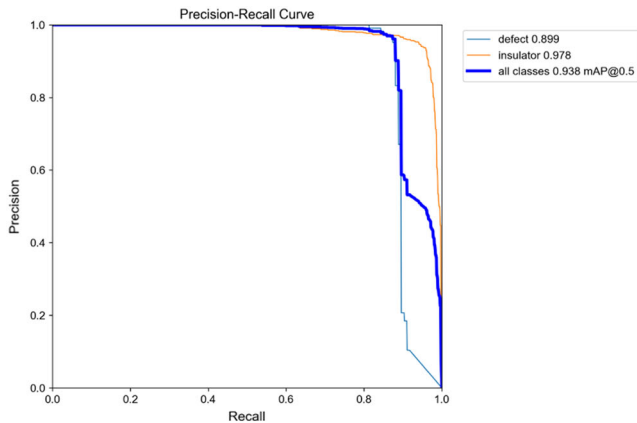
(a) YOLOv5n improved pre-P-R curve
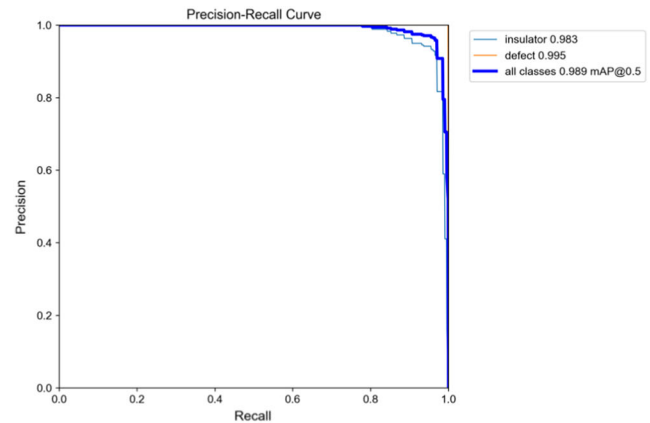
(b) YOLOv5n Modified curve

Fig. 10 P–R curve before and after improvement

4.3.5 Comparison experiments of different detection algorithms

By comparing the performance of the improved lightweight detection model YOLOv5n with other mainstream detection algorithms and the existing improved lightweight YOLOv5 algorithm, the data in this study are the results of 300 rounds of experimental epochs, the specific results are shown in Table 8, and the P–R curve before and after improvement is shown in Fig. 10.

5 Conclusion

This study is based on the original YOLOv5n algorithm. By replacing the C3 module with the C2f module of YOLOv8 on the basis of the original YOLOv5n, the SimAM attention mechanism is introduced, and the loss function is improved to the Wise-IoU loss function to improve the detection of insulator defects on the transmission line. Compared with YOLOv5, YOLOv8 is more lightweight, and the improvement and comparison experiments in the above three parts prove that the improved algorithm has significantly improved the accuracy and recall rate in defect detection on the public data set and custom data set. It shows the feasibility of

the improved research and can be better applied to similar complex and variable image detection environment.

Author Contributions Yachao Si provides financial support, research concepts, paper review and revision; Jiajie Gao is responsible for data collation and management, writing the first draft of the paper, collecting and analyzing experimental data in Tables 1, 2, 3, 6 and 7, and verifying the experimental design; Mingzhan Zhao provided experimental equipment, graphics cards and servers, collected and modified the content of Fig. 1, and improved and displayed Figs. 2, 3, 4 and 6; Xiaojun Xu collected and reviewed the references and analyzed the research results and experimental data. Manuscripts were reviewed by all authors.

Declarations

References

1. Chao, J., Xinbo, H., Wen, C., et al.: Research on improved Fast-CNN model in insulator feature detection. Computer & Modernization, (04):59-64+71 (2019)
2. Haoyang, C., Yi, H., Xun, Z., et al.: Motion blurred insulator image restoration method based on improved U-net structure genera-

tive adversarial network. Power Grid Technol. **47**(06), 2594–2604 (2023)

3. Zhang, X., An, J., Chen, F.: A method of insulator fault detection from airborne images. In *Wri Global Congress on Intelligent Systems*. IEEE Computer Society, 200-203 (2010)

4. Jucai, L., Jun, H., Fangming, C., et al.: Defect diagnosis of glass insulators based on color images. Power Grid Technol. **35**(01), 127–133 (2011)

5. Kaijun, W., Zehao, X., Hongquan, S.: Rapid detection method for self-explosion defects of Glass insulators based on improved FasterNet and YOLOv5. High Voltage Technol., 1-11

6. Ying, Z.Y.: Research on insulator detection technology based on deep learning. Harbin Engineering University, (2022)

7. Mei, Zhang, Zhijia, Qi., Panxia, He.: Insulator defect detection algorithm based on improved YOLOv5. J. Anhui University Sci. Technol. **43**(04), 10–18 (2023)

8. Zhenliao, W., Zhengping, W., Shuifa, S.: Insulator multi-defect detection based on improved YOLOv5 algorithm. High Voltage Apparatus, 1-11

9. Niantao, Wang, Shuqing, Wang, Tang, Lu., Dan, Ma.: Insulator defect detection based on EfficientNet-YOLOv5s. J. Hubei Univ. Technol. **38**(01), 21–26 (2023)

10. Deng, F. et al.: A novel insulator defect detection scheme based on deep convolutional auto-encoder for small negative samples. High Voltage (2022)

11. Linghua, Q., Zhengtao, Z.: Research on Insulator defect detection of transmission line based on deep learning. Appl. Res. Comput. S01(2020):4

12. Yang, L. et al.:SimAM: A Simple, parameter-free attention module for convolutional neural networks. Int. Conf. Mach. Learn. (2021)

13. Tong, Z. et al.: Wise-IoU: bounding box regression loss with dynamic focusing mechanism. arXiv:2301.10051 (2023)

14. Wang, C.-Y. et al.: CSPNet: A New Backbone that can Enhance Learning Capability of CNN. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (2019): 1571-1580