Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105 (Online)
Volume **8**, Number **3**, August **2023**

Saliency Improved YOLOv5s Algorithm for Defect Detection of Overhead Line Insulator

Xiang Fang

School of Electrical Engineering
Henan University of Technology
Zhengzhou, China
xiangfang@stu.haut.edu.cn

Xiang Wu*

School of Electrical Engineering
Henan University of Technology
Zhengzhou, China
xiangw@haut.edu.cn

Yuanhao Ma

School of Electrical Engineering
Henan University of Technology
Zhengzhou, China
mayuanhao@stu.haut.edu.cn

Hao Lian

School of Electrical Engineering
Henan University of Technology
Zhengzhou, China
lianhao@stu.haut.edu.cn

Yuan Song

Istituto della Ricerca ed Innovazione Sino-Italian
Marche Polytechnic University
Marche, Italy
istituto.macsi@gmail.com

*Corresponding author: Xiang Wu

Received January 13, 2023, revised March 9, 2023, accepted May 25, 2023.

Abstract. *Aiming at the problems that the target proportion of insulator defects is small, and the defect features are not obvious, which leading to low detection rate, a YOLOv5s algorithm based on saliency information improvement is proposed. Firstly, saliency information extraction (SANet) module is used to generate saliency images to enhance the discrimination between insulator target and background. Second, a small target detection layers are added, and a large target detection layer are removed to improve the feature extraction ability of the network for small targets and speed up the network reasoning. Then, the attention module (CBAM) is introduced to improve the network's attention to the target features along both channel and spatial dimensions. Finally, EIoU-Loss as a loss function is used to accelerate the convergence speed and improve the regression accuracy. The experimental results show that compared with the original YOLOv5s network, the improved algorithm with mAP50 and mAP50:95 increases by 3.43% and 3.32%, respect to the frame rate 36 frames per second, which is capable of real-time detection.*

Keywords: Insulator defects, YOLOv5s, Saliency information, CBAM, Target detection

1. Introduction.

1.1. Background.

Insulators are the key components of high-voltage transmission lines, which work in outdoor environment for a long time and their surfaces are highly susceptible to insulator breakdown and damage due to moisture erosion and pollution ash accumulation [1], so it is necessary to inspect insulators of transmission lines regularly by means of UAVs and other means. Since the insulator defects in the collected images have small percentage, inconspicuous features, and complex background, which seriously affect the recognition and detection accuracy of insulator defects. Therefore, a detection method for small target defects in complex scenes is urgently needed to improve the recognition and detection accuracy of insulator defects.

In recent years, with the continuous expansion of artificial intelligence applications [2], deep learning-based target detection algorithms [3] have seen a spurt in development. Compared with traditional detection methods, deep learning-based detection methods can fully utilize the computing power of computers and learn target features autonomously through input data, solving the problems of small number of features, poor robustness, and low recognition rate in previous traditional methods, and have better adaptation and extraction capabilities for target features in different scenes [4].

The current insulator defect algorithms based on deep learning are divided into two categories, one is two-stage detection algorithm, such as literature [5] proposed Faster-RCNN algorithm based on dynamic selection network improvement to achieve the detection of insulators; literature [6] used Cascade-RCNN network as the basic framework to train multiple cascaded detectors to complete the detection of insulator defects by setting different thresholds detection. The above two-stage detection algorithms represented by Faster-RCNN[7] and Cascade-RCNN[8] have relatively high detection accuracy, but the detection speed is slow. Another category is one-stage detection algorithms, such as the MobileNet-SSD-based overhead line defect detection algorithm in literature [9]; literature [10, 11] improves the detection accuracy of insulator defects based on YOLOv4 with the loss function of a priori anchor frame and improved regression border; the YOLOv5 model is improved by incorporating attention mechanism in literature [12], which effectively improves the YOLOv5 model is improved by incorporating attention mechanism in the literature, which effectively improves the focusing ability of the target to be inspected and achieves high accuracy detection of insulator defects in complex scenes. The YOLO series algorithm represented by YOLOv5 [13] stands out in the first stage of detection

algorithms with its good detection accuracy and fast detection speed, and is widely used in industrial inspection, but the algorithm is mainly for the detection of targets of medium size. Considering that insulator defects have a small target share and inconspicuous target features, it is still challenging for YOLOv5s algorithm to achieve accurate detection of small target defects in complex backgrounds.

To address the above problems, this paper proposes an improved YOLOv5s insulator detection algorithm based on saliency information using the YOLOv5s network as a framework, generating saliency images through the saliency information extraction (SANet) module to improve the extraction of target features in complex backgrounds; adding the detection of smaller-scale targets to the Head layer of the original YOLOv5s network, removing The detection of large scale targets is removed from the Head layer of the original YOLOv5s network, which improves the detection accuracy of smaller targets and reduces the complexity of the network; a lightweight attention module (CBAM)[14] is introduced to reduce the attention of the network to invalid information and reduce the false detection rate of insulator defects; the EIoU[15] loss function is used to replace the regression edge (CIoU) loss function to further improve the model performance. Experiments prove that the improved model has a significant improvement in detection accuracy compared with the original YOLOv5s network and can meet the minimum requirement of 24 seconds/frame for real-time detection with real-time detection capability.

1.2. **YOLOv5s network.** YOLOv1, as the pioneer of the first stage of target detection algorithms, makes the YOLO series algorithms have a very important position in the field, and from YOLOv1 to YOLOv5, the YOLO series algorithms have been gradually optimized and achieved wide popularity in the industrial detection field. yOLOv5s is a version of YOLOv5, which considers both detection accuracy and detection speed, and the network inference is faster, and the computational complexity is lower, which is easier to deploy in embedded systems compared to mainstream target detection algorithms such as SSD [16], Cascade-RCNN, and YOLO series, so this paper is based on YOLOv5s network for improvement.

The structure of YOLOv5s network is shown in Figure 1, which consists of three parts: feature extraction (Backbone) layer, feature fusion (Neck) layer and feature detection (Head) layer. the Backbone layer of YOLOv5s network uses cross-stage local (CSP) network as the backbone network, including five standard convolutional modules, four C3 modules and one spatial pyramid pooling (SPP) module. Among them, the C3 module is used for the learning of residual features and consists of three convolutional layers as well as multiple Bottleneck, which effectively reduces the number of model parameters and computational complexity, and the SPPF module is used for feature integration, which speeds up the feature fusion without affecting the accuracy; the Neck layer of YOLOv5s network draws on the YOLO series algorithm Feature Pyramid (FPN) and Path Aggregation Network (PAN) interlinked feature fusion, where FPN predicts to large scale feature maps by up-sampling in a bottom-up manner to enhance image semantic features, and PAN predicts to small scale feature maps by down-sampling in a top-down manner to enhance image localization features, aggregating the features generated by FPN and PAN at different scales of the detection layer. The Head layer of YOLOv5s accomplishes target classification and location prediction on the feature maps at three scales of 8x, 16x, and 32x of the input original image, respectively.

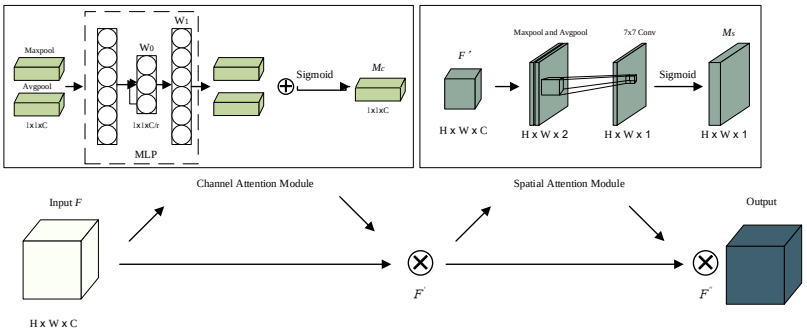


FIGURE 1. Structure of YOLOv5s network improved based on saliency information

2. **Improved YOLOv5s network based on saliency information.** For the problems of low recognition accuracy and poor localization accuracy in YOLOv5s network for small target detection, this paper improves YOLOv5s network, and the specific improvement work is as follows: 1) add the Significance Information Extraction (SANet) module, extract the significant target region by semantic segmentation network, generate the significant image and then integrate it into Neck layer, which effectively suppresses the background 2) add up-sampling operation in the feature pyramid network (FPN) to get higher resolution feature maps, and add 160*160 smaller target detection and remove 20*20 large target detection in the Head layer to improve the small target feature expression ability while reducing the model complexity; 3) introduce attention (CBAM) module to improve the network for useful information capture; 4) introduce EIoU loss function instead of CIoU loss function to improve the regression accuracy of the model.

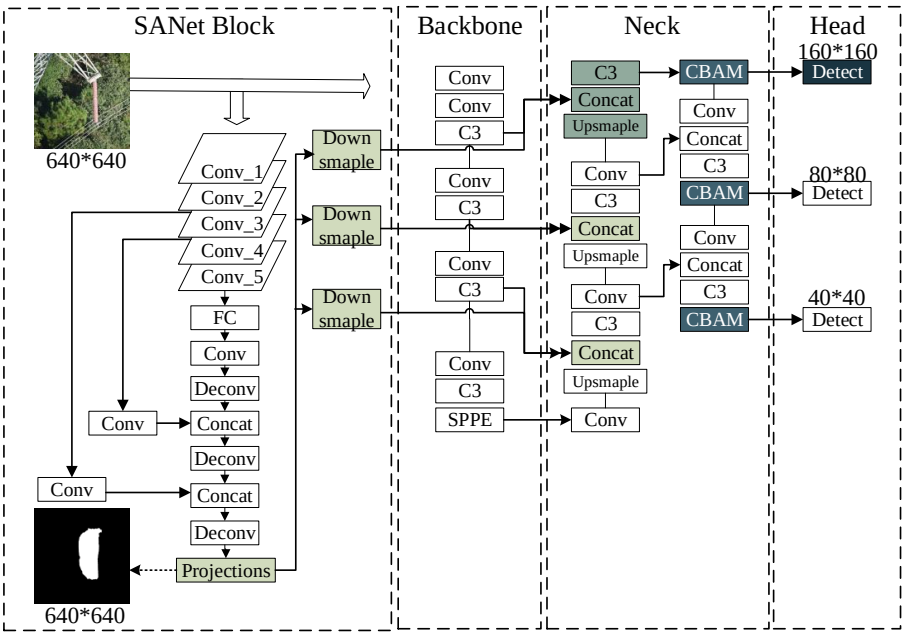


FIGURE 2. Structure of YOLOv5s network improved based on saliency information

2.1. **Significance Information Extraction (SANet) Module.** The extraction of saliency information is essentially a segmentation task, which is accomplished by identifying the salient subjects of an image and then performing pixel-level segmentation and is widely used in many fields such as robot collaboration, target detection, and vision tracking. In the task of insulator defect detection, the interference of complex background makes the extraction of insulator defect features more difficult. Therefore, SANet network sets insulators as saliency targets, suppresses the influence of background information on insulator defect detection by enhancing the extraction of insulator features, and improves the detection accuracy of the network for insulator defects.

In order to extract the insulator features in the image, this paper refers to literature [17] to design SANet network to separate insulator information from background information, the network structure is shown in Figure 3, where the existence of rich edge information in *conv*_3 layer helps to improve the accuracy of target localization, and the convolutional features in *conv*_4 layer and *conv*_5 layer help to improve the recognition accuracy of the target, so the Therefore, the feature maps generated by *conv*_3,*conv*_4 and *conv*_5 are fused to improve the localization accuracy and recognition accuracy of salient targets in SANet networks, and the specific process is as follows: input the original image, add a 1×1 convolutional block to *conv*_3_3, *conv*_4_3 and *conv*_5_3 feature maps after two fully connected layers for pixel-by-pixel category prediction, denoted as *conv*_5 prediction, *conv*_4 prediction and *conv*_3 prediction. The feature map predicted by *conv*_5 is cascaded with the feature map predicted by *conv*_4 through the deconvolution up-sampling operation to generate a feature map with more deep semantic information, and to obtain better target edge information, it is fused with the feature map predicted by *conv*_3, and the insulator prediction map of the original image size is obtained through the deconvolution up-sampling operation with a convolution kernel of 4×4.

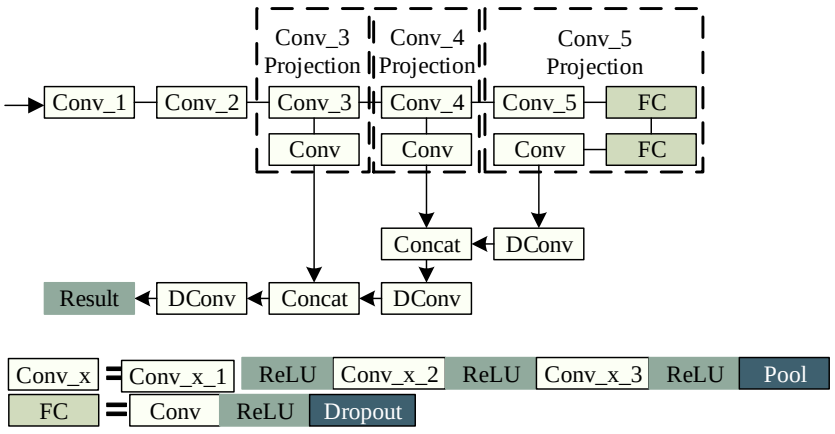
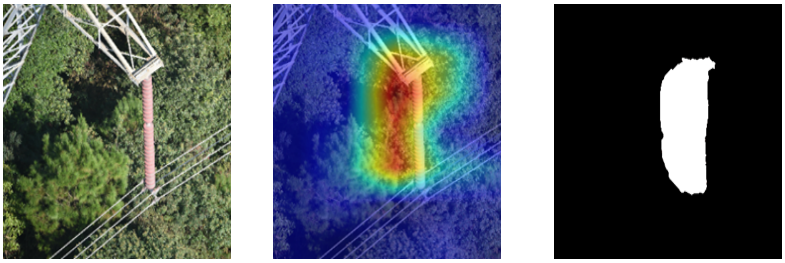


FIGURE 3. SANet network structure

The insulator target is predicted pixel by pixel using SANet network, and the predicted insulator part is labeled as foreground and the rest is labeled as background to output the saliency image prediction results. The original image, the visualization results of saliency features and the prediction results of saliency images are shown from left to right in Figure 4. From the results in Figure 4(b) and Figure 4(c), the SANet network can extract the insulator information intact. Since the task of saliency images is only used to assist detection and provide saliency information for the Neck layer of YOLOv5s, only

coarse localization of saliency information and segmentation of approximate contours are needed to help the YOLOv5s network enhance the discrimination of insulator features from background features to improve the localization accuracy of insulator defects.

The YOLOv5s network uses the feature fusion of FPN+PAN to generate feature maps for detection at three scales: large, small, medium, and large. To make full use of the saliency information, the saliency images are down-sampled at multiple levels and incorporated into the feature maps at different scales in the YOLOv5s network to assist the network feature detection layer in enhancing the localization and extraction of saliency features.



(a) Original image (b) Visualization result (c) Prediction result

FIGURE 4. Significance graphs generated based on SANet

2.2. **SANet module loss function.** The loss function of SANet network consists of three parts: classification loss, regression loss and segmentation loss, in which cross-entropy loss function and smooth loss function are used for classification loss and regression loss, respectively. In this paper, the segmentation separates the saliency information in the image from the background, which is a dichotomous auxiliary detection task, so the same loss function is used for the segmentation loss and classification loss in this paper.

$$L_{SANet}\{\mathrm{c,t,m}\} = \alpha \sum_i L_{cls}(c_i, c_i^*) + \beta \sum_i L_{reg}(t_i, t_i^*) + \sum_i L_{mask}(m_i, m_i^*) \tag{1}$$

where α is the weight coefficient of classification loss in SANet network and β is the weight coefficient of regression loss. α and β take the values of 0.2 and 0.8 respectively in the model training.

1) Classified loss L_{cls}

$$\sum_i L_{cls}(c_i, c_i^*) = -[c_i ln c_i^* + (1 - c_i) \ln (1 - c_i^*)] \tag{2}$$

where $\mathcal{C}_i$ denotes the probability that the feature in the $i-$th prediction frame is an insulator, and $\mathcal{C}_i^*$ denotes the probability that the intersection ratio of the $i-$th real frame and the prediction frame is an insulator, if the intersection ratio of the $i-$th real frame and the prediction frame is greater than 0.5, then the prediction result is a positive sample $\mathcal{C}_i^* = 1$, and vice versa is a negative sample $\mathcal{C}_i^* = 0$.

2) Regression loss L_{reg}

$$\sum_i L_{reg}(t_i, t_i^*) = \sum_{j \in x,y,w,h} Smooth_{L1}(t_{i,j} - t_{i,j}^*) \tag{3}$$

where j denotes a one-dimensional vector containing coordinate and aspect information, $t_{i,j}$ denotes the coordinate and aspect information of the $i-$th prediction frame, denotes

the coordinate and aspect information of the $i-$th true frame, and the SmoothL1 loss function is defined as

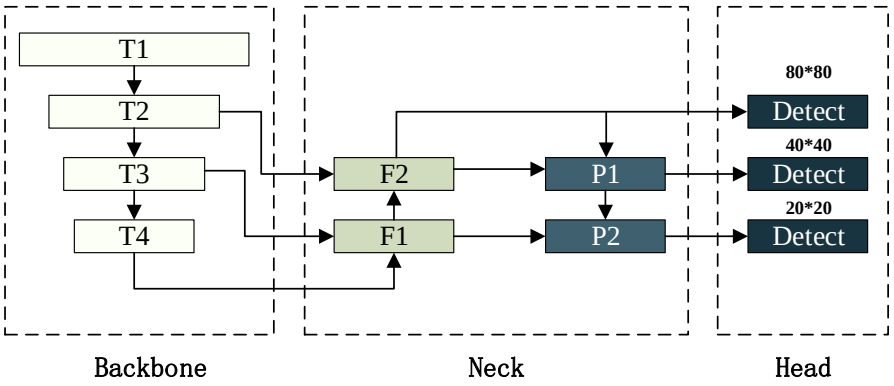
$$Smooth_{L1}\begin{cases}0.5x^2, |x|<1\\ |x|-0.5, other\end{cases} \tag{4}$$

3) Segmentation loss L_{mask}

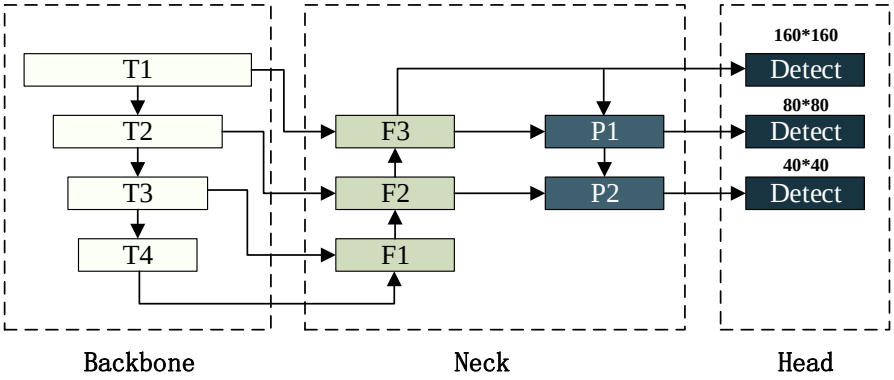
$$\sum_i L_{mask}(m_i, m_i^*) = -[m_i lnm_i^* + (1-m_i)\ln(1-m_i^*)] \tag{5}$$

Where, m_i denotes the probability that the $i-$th mask is an insulator, $m_i{}^*$ denotes the $i-$th mask frame true frame label, target region $m_i{}^* = 1$, and conversely, the background region $m_i{}^* = 0$.

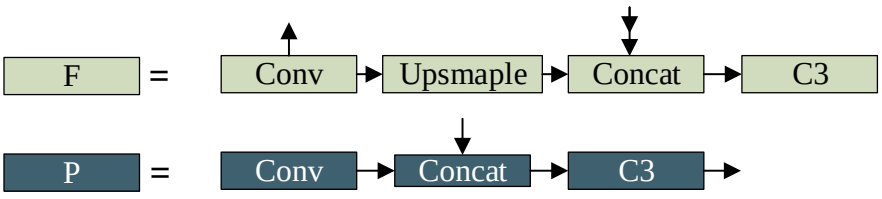
2.3. **Improvement of multi-scale detection layer.** As shown in Figure 5(a), the Head layer of YOLOv5s network has three scales of detection layers to complete the target classification and position prediction on the feature maps of 8x, 16x and 32x scales of the input original image, i.e., when the input size is 640*640, T2, T3 and T4 extract 80*80, 40*40 and 20*20 feature maps respectively, which are fed into the Head layer for large, medium and small target detection after multi-scale feature fusion in the F and P layers, respectively.



(a) YOLOv5s multi-scale fusion and detection map



(b) Improved multi-scale fusion and detection map



(c) Multi-scale feature fusion and detection improvement diagram

FIGURE 5. Significance graphs generated based on SANet

UAV inspection is the usual way of insulator fault detection, and the large field-of-view property [18] of UAV makes insulator defects occupy a relatively small portion of the image, which makes it difficult to complete accurate detection of insulator defects with the 80*80 size feature map used in YOLOv5s network. In view of the feature that the features of small targets on the high-resolution feature map are more expressive and easier to be detected, this paper improves the multi-scale detection layer of YOLOv5s network, and the improved network is shown in Figure 5(b), and the specific work is as follows: 1) After continuing the up-sampling operation on the feature map generated by the F2 layer, it is fused with the T1 layer features to obtain a higher resolution feature map, adding

160*160 scale detection head to improve the detection accuracy of the detection layer for smaller targets; 2) most of the target defect features account for small and medium sizes, so the large target detection layer of 20*20 scale in the original YOLOv5s network is removed to reduce the model computational complexity and improve the operation speed.

2.4. **CBAM module.** The literature [19] in 2018 proposed a lightweight and efficient attention module, which inferred the attention mapping graph along two dimensions, channel and space, where the channel attention mechanism calculates the weight coefficients of each channel of the input image through the network, and the spatial attention mechanism finds the part with the most aggregated location information based on the direction of the channel based on the channel attention to obtain the channel-space attention graph that improves the feature representation of the network for the target region. Due to the complex environment in the insulator detection images used in the dataset and the small percentage of insulator defect parts compared to the image pixels, to achieve accurate detection of insulator defects, this paper embeds the CBAM attention module into the Neck layer based on the improved multi-scale detection layer, and the network structure is shown in Figure 6.

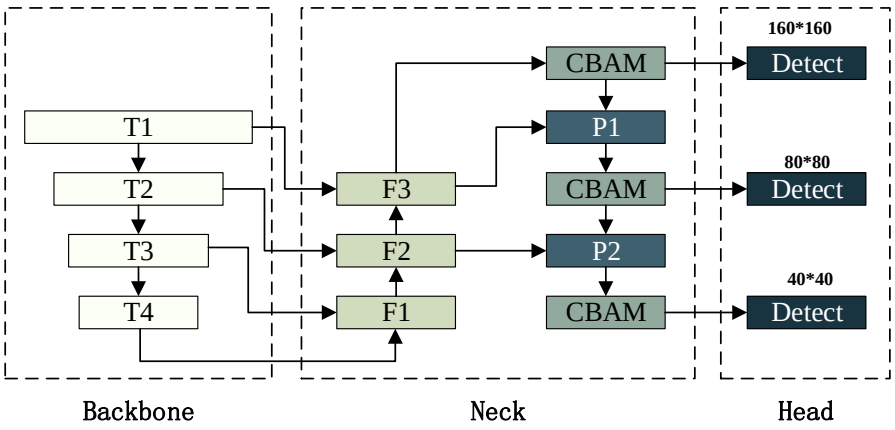


FIGURE 6. Schematic diagram of YOLOv5s+CBAM network structure

The specific process of CBAM module is as follows, first set the middle layer feature map $F \in \mathbb{R}^{C\times H\times W_*}$ as input, send it to Channel Attention Module (CAM), keep the input feature map channel dimension unchanged, compress the spatial dimension to generate the attention map $M_c(F) \in \mathbb{R}^{C\times 1\times 1}$, multiply with the input feature residuals to get the channel adaptive feature refinement map , then through the Spartial Attention Module (SAM) module, use the spatial dimension of the feature map to get the spatial attention map, and finally output the adaptive feature map after weighted multiplication. Spartial Attention Module (SAM) module, using the spatial dimension of the feature map to obtain the spatial attention map $M_s(F) \in \mathbb{R}^{1\times H\times W}$, and finally output the adaptive feature map F'' after weighted multiplication. The attention calculation formula is shown in Equation (6)(7), and the structure of CBAM is shown in Figure 7:

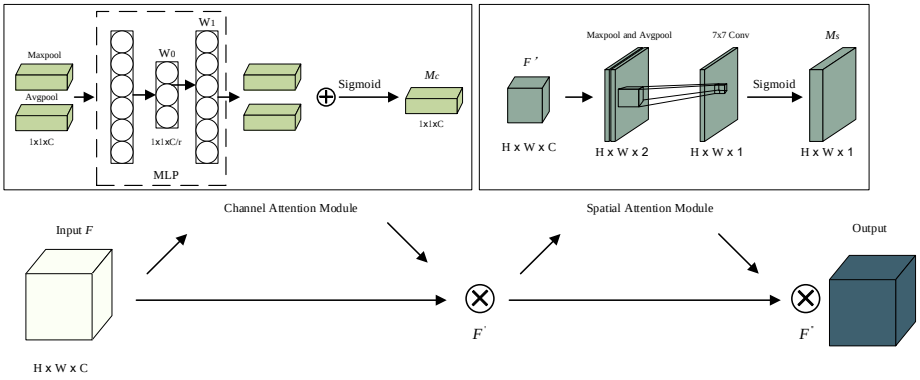


FIGURE 7. CBAM Module schematic

$$F' = M_C(F) \otimes F \tag{6}$$

$$F'' = M_S(F) \otimes F' \tag{7}$$

2.5. **Loss function.** The loss function of YOLOv5 network consists of three parts: bounding box loss, classification loss, and confidence loss. CIoU-Loss, Logits, and Binary Cross Entropy loss functions are used to calculate the target bounding box, classification, and confidence losses, respectively. At present, the most widely used loss function in the bounding box regression is CIoU-Loss, but in its regression process, there is a problem that the width and height of the prediction box cannot be increased and decreased at the same time when the w and h of the prediction box and the real box present a linear ratio. To address the above problems, EIoU is improved since CIoU loss function by introducing the difference between the width and height of the prediction frame and the width and height of the minimum external matrix and redefining the aspect ratio loss term to make the prediction frame more consistent with the real frame. Therefore, EIoU-Loss is finally chosen to replace CIoU-Loss as the loss function of the regression frame in this paper, and the EIoU calculation process is as follows:

$$L_{EIOU} = L_{IOU} + L_{dis} + L_{asp} \tag{8}$$

$$L_{IOU} = 1 - IOU \tag{9}$$

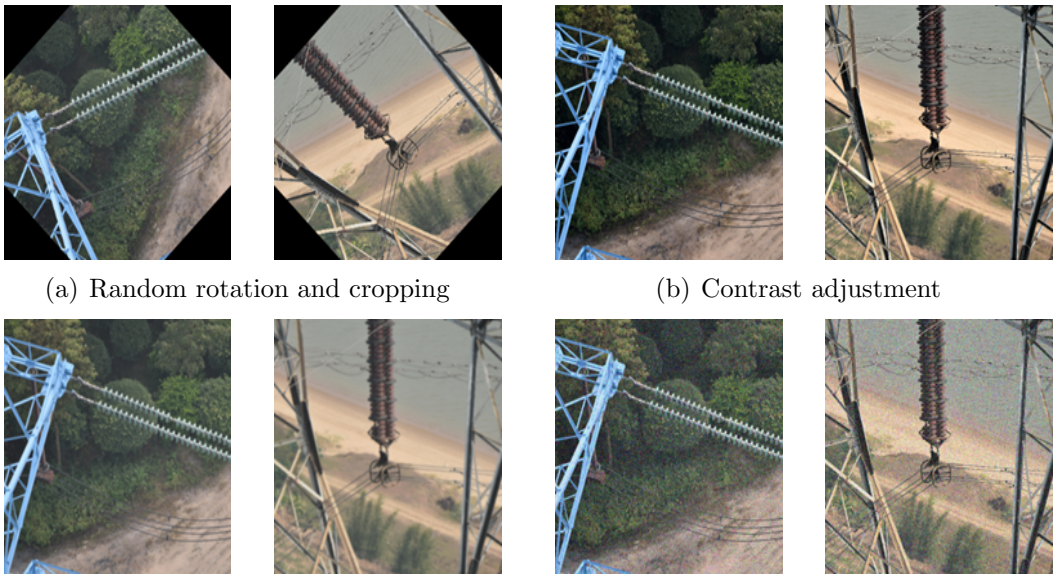
$$L_{dis} = \frac{\rho^2(b, b^{gt})}{c^2} \tag{10}$$

$$L_{asp} = \frac{\rho^2(w, w^{gt})}{c_w^2} + \frac{\rho^2(h, h^{gt})}{c_h^2} \tag{11}$$

where IOU (Intersection of Union) is the intersection and ratio of the prediction frame and detection frame, which indicates the degree of overlap between our detection area and the target area, and it takes values in the range of [0,1], L_{IOU} denotes the overlap loss, L_{dis} denotes the center distance loss, L_{asp} denotes the aspect ratio loss, $\rho^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$ denotes the Euclidean distance between two points, c denotes the diagonal distance of the minimum external matrix of the overlap, w , h , and b denote the width, height, and center point coordinates of the prediction frame, respectively, w^{gt} , h^{gt} , and b^{gt} denote the width, height, and center point coordinates of the real frame, respectively, C_w and C_h denote the width and height of the direct minimum external matrix of the prediction frame and the real frame.

3. Experiments and results analysis.

3.1. **Experimental data set.** The experimental dataset consists of insulator images collected by China Power Grid, and there are 3409 insulator images in the dataset, including 892 images of insulator defects. Therefore, this paper adopts random rotation and cropping, contrast adjustment, Gaussian blur, and noise injection to expand the insulator defect data set, and some of the image enhancement effects are shown in Figure 8. In this paper, 3568 insulator defect images are randomly divided into training set, test set and validation set in the ratio of 8:1:1, and then the insulator defects are manually labeled by via toolbox, and the labeled images are saved in VC2007 format for model training.



(a) Random rotation and cropping (b) Contrast adjustment

(c) Gaussian blur (d) Injecting noise

FIGURE 8. Example of partial image enhancement

3.2. **Descriptive statistics.** The experimental configuration of this paper is: CPU: Intel(R) Core(TM) i5-1135G7, GPU: NVIDIA RTX 1080Ti, OS: Windows 10, deep learning framework: Pytorch 1.7.0, python 3.8, CUDA 11.4. the hyperparameter settings of the improved model are shown in Table 1.

TABLE 1. Model hyperparameter settings

Learning rate	Number of iterations	Batch size	Decay weights	Optimizer
0.01	300	32	0.0005	SGD

To verify that the improved algorithm proposed in this paper has superiority in detection accuracy, the widely recognized correct rate P(%), recall rate R(%) and average accuracy mAP(%) in the field of target detection are chosen as indicators to evaluate the model performance, as follows:

$$P = \frac{TP}{TP + FP} \tag{12}$$

$$R = \frac{TP}{TP + FN} \tag{13}$$

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

$$mAP = \frac{1}{n} \sum_{i=0}^{n} AP(i) \tag{15}$$

where TP is the number of insulator defect samples with correct prediction results, FP is the number of insulator defect samples with incorrect prediction results, and FN is the number of non-insulator defect samples with incorrect prediction results.

3.3. **Ablation experiments.** To verify the effectiveness of the improved algorithm in this paper, ablation experiments are conducted for each module, and the experiments compare (1) YOLOv5s model; (2-5) YOLOv5s +EIoU model, YOLOv5s +CBAM model, YOLOv5s +improved detection layer model and YOLOv5s based on YOLOv5s by adding improvement mechanisms one by one +SANet model; (6) YOLOv5s +SANet +improved detection layer model based on significance information extraction and multi-scale detection layer improvement; (7) YOLOv5s +SANet +improved detection layer +CBAM model based on serial number 6 with the introduction of CBAM module in front of different scale detection layers; (8) EIoU based on serial number 7 with loss function to replace the CIoU loss function in the original model for the improved model in this paper. In this paper, the above eight models are experimented on insulator defect detection dataset using P, mAP_{5o} and $mAP_{5o:95}$ as evaluation indexes, Where mAP_{5o} indicates the detection accuracy when the IOU cross-merge ratio threshold is 0.5, when the confidence of the predicted sample is greater than the threshold, the sample is predicted to be positive, and vice versa. $mAP_{5o:95}$ indicates the average detection accuracy of 10 different cross-merge ratios in 50-95, and the experimental results are shown in Table 2.

Table 2. Ablation experiments

NO.	SANet	Detection layer	CBAM	EIoU	P(%)	mAP_{5o}(%)	$mAP_{5o:95}$(%)
1					89.54	86.75	53.17
2				✓	90.26	86.94	53.41
3			✓		90.67	87.05	53.34
4		✓			91.28	88.29	54.88
5	✓				92.12	89.11	55.63
6	✓	✓			93.23	86.97	56.41
7	✓	✓	✓		93.55	89.92	56.36
8	✓	✓	✓	✓	93.79	90.18	56.49

As can be seen from the experimental results in Table 2, Sequence No. 2 uses EIoU to replace the loss function of the original network, which improves the regression accuracy of the prediction frame and improves P by 0.72% compared to the YOLOv5s network; Sequence No. 3 adds the CBAM module to the original network and improves P by 1.13% compared to the YOLOv5s network; Sequence No. 4 adds a smaller target detection layer to the multiscale detection layer to improve The detection accuracy of the network for small targets is improved by 1.54 and 1.71% compared with YOLOv5s network mAP_{5o} and $mAP_{5o:95}$, respectively; the addition of SANet module to the original network by serial number 5 improves mAP_{5o} and $mAP_{5o:95}$ by 2.36% and 2.46%, respectively. positive effect.

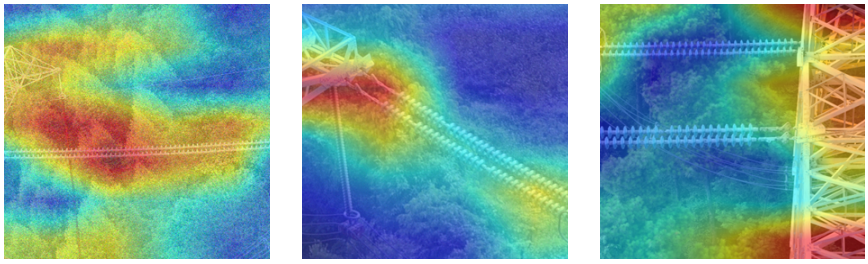
Finally, Sequence 8 improves the multi-scale detection layer by fusing both SANet saliency information module and CBAM attention module into the network, and replaces

the original network loss function with EIoU loss function, which effectively improves the detection accuracy of the network, compared with the original YOLOv5s network, the detection accuracy P improves by 4.25%, mAP_{5o} improves by 3.43%, and $mAP_{5o:95}$ improves by 3.32%. The detection accuracy P is improved by 4.25%, mAP_{5o} by 3.43% and $mAP_{5o:95}$ by 3.32% compared with the original YOLOv5s network. By combining the above improvements, the detection accuracy P can reach 93.79%, mAP_{5o} and $mAP_{5o:95}$ can reach 90.18% and 56.49% respectively, which has a better detection effect in the detection of insulator defects in complex scenarios.

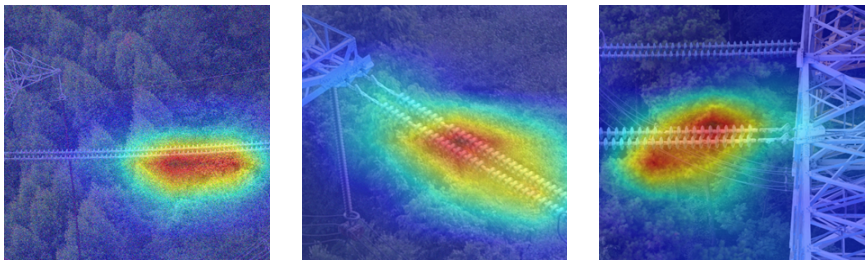
To further verify the effectiveness of the improved algorithm, some images from the insulator defect dataset are selected for testing, Figure 9(b)(c) shows the feature visualization results of the YOLOv5s algorithm and the algorithm in this paper for insulator defect detection, respectively, from the experimental results, it can be seen that the features of insulator defects are not obvious in the YOLOv5s algorithm network, while the improved algorithm in this paper can effectively suppress the influence of background information on detection and improve the network's attention to the features of small target insulator defects. Figure 10 shows the detection effect of insulator defects, and the improved YOLOv5s algorithm based on saliency information proposed in this paper achieves the accurate detection of small target insulator defects in complex environments, which has some practical value.

Figure 11 shows the comparison of $mAP_{5o}(\%)$ between this algorithm and the original algorithm YOLOv5s in 300 training rounds. From the figure, it can be seen that the improved model gradually converges around the 55-th round, and the original algorithm YOLOv5s gradually converges around the 100-th round, compared with the original algorithm, this algorithm has obvious improvement in convergence speed and accuracy.

(a) Original image



(b) YOLOv5s algorithm feature visualization results



(c) The visualization results of the algorithm features in this paper

FIGURE 9. Comparison of insulator feature visualization results

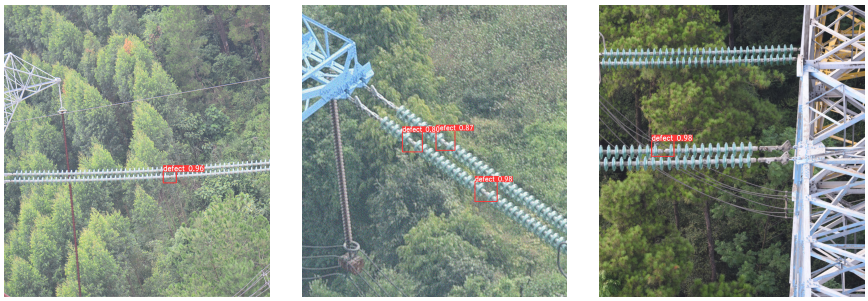


FIGURE 10. Insulator defect detection results comparison

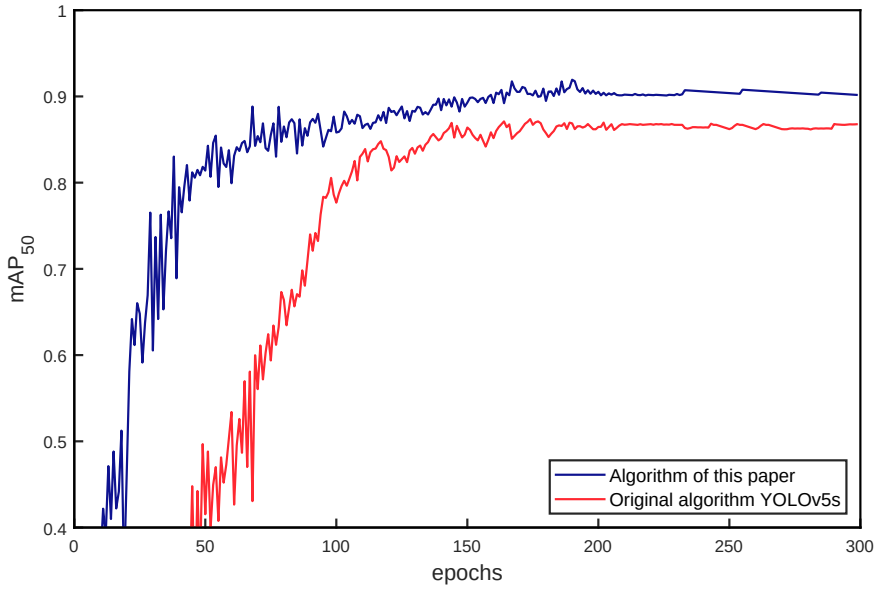


FIGURE 11. Comparison chart of loss function training.

3.4. **Comparison experiments.** To verify the superiority of the algorithm in this paper, mAP_{5o} and FPS (frames per second) are chosen as evaluation indexes to compare this algorithm with mainstream target detection algorithms, and the comparison results are shown in the following table.

TABLE 3. Comparison of mainstream target detection algorithms.

N0.	Algorithm name	mAP_{5o}(%)	FPS(Hz)

1	SSD	75.33	26
2	Cascade-RCNN	81.97	19
3	Faster-RCNN	84.65	16
4	YOLOv4	87.37	43
5	YOLOv5s	86.75	107
6	YOLOR	88.23	33
7	Our	90.18	36

As can be seen from Table 3, compared with SSD, Cascade-RCNN, and Faster-RCNN algorithms, this paper has significant advantages in detection accuracy and speed, and compared with YOLO series target detection algorithms, this paper has improved the detection accuracy by 2.57%, 3.43% and 1.95% respectively. Although the detection speed is slightly inferior, it still has the ability of real-time detection. In summary, this algorithm has higher detection accuracy than the current mainstream insulator defect detection algorithms based on real-time detection requirements.

4. **Conclusion.** In this paper, we propose an improved YOLOv5s algorithm based on saliency information to address the problem of low detection rate of insulator defect recognition in complex scenes, generating saliency images through SANet network to suppress the influence of background information; improving the multiscale detection layer to enhance the detection capability of the network for smaller targets; introducing CBAM module in Neck layer to improve the network for defect features The CBAM module is introduced in the Neck layer to improve the focus of the network on the defect features; the EIoU-Loss is replaced to accelerate the convergence of the model and improve the regression accuracy. The average detection accuracy of mAP_{5o} and $mAP_{5o:95}$ reaches 90.18% and 56.49%, respectively, which is an improvement of 3.43% and 3.32%, and effectively reduces the false detection and leakage of insulator defects, and the algorithm has better recognition and performance than other target detection algorithms.

REFERENCES

[1] J. Xu, Q. Wang, "Analysis of Natural Pollution Deposit on Optic Sensor for Monitoring of insulator Contamination," *Applied Mechanics and Materials*, vol. 721, 424–427, 2015.

[2] C.-M. Chen, S. Lv, J. Ning, and J. Wu, "A Genetic Algorithm for the Waitable Time-Varying Multi-Depot Green Vehicle Routing Problem," *Symmetry*, vol. 15, 1, 124, 2023.

[3] Y.-L. Zhao, Y.-Q. Wu, "Research progress of surface defect detection methods based on machine vision," *Journal of Scientific Instruments*, vol. 43, 198–219, 2022.

[4] Z. Xie, J. Chen, Y. Feng, K. Zhang, Z. Zhou, "End to end multi-task learning with attention for multi-objective fault diagnosis under small sample," *Journal of Manufacturing Systems*, vol. 62, 301–316, 2022.

[5] Y. Ma, Y. Zhang, "Insulator detection algorithm based on improved Faster-RCNN," *Journal of Computer Applications*, vol. 42, 2, 631, 2022.

[6] Q. Wen, Z. Luo, R. Chen, "Deep learning approaches on defect detection in high resolution aerial images of insulators," *Sensors*, vol. 21, 4, 1033, 2021.

[7] W. Zhao, M. Xu, X. Cheng, Z. Zhao "An insulator in transmission lines recognition and fault detection model based on improved faster RCNN," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, 1–8, 2021.

[8] J. Han, Z. Yang, H. Xu, G. Hu, C. Zhang, H. Li, S. Lai, H. Zeng "Search like an eagle: A cascaded model for insulator missing faults detection in aerial images," *Energies*, vol. 13, 3, 713, 2020.

[9] Y. Yang, Y. Wang, H. Jiao "Insulator identification and self-shattering detection based on mask region with convolutional neural network," *Journal of Electronic Imaging*, vol. 28, 5, 053011–053011, 2019.

[10] L. Yang, J. Fan, S. Song, Y. Liu "A light defect detection algorithm of power insulators from aerial images for power inspection," *Neural Computing and Applications*, vol. 34, 20, 17951–17961, 2022.
[11] Z. Qiu, X. Zhu, C. Liao, D. Shi, W. Qu "Detection of transmission line insulator defects based on an improved lightweight YOLOv4 model," *Applied Sciences*, vol. 12, 3, 1207, 2022.
[12] J. Du, Y. Jiang "Improved YOLOv5-based Method for Recognition of Insulator String Pins," *2022 12th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, 882–887, 2022.
[13] Z. Feng, L. Guo, D. Huang, R. Li "Electrical insulator defects detection method based on yolov5," *2021 IEEE 10th Data Driven Control and Learning Systems Conference (DDCLS)*, 979–984, 2021.
[14] L. Wang, Y. Cao, S. Wang, X. Song, S. Zhang, J. Zhang, J. Niu "Investigation into recognition algorithm of Helmet violation based on YOLOv5-CBAM-DCN," *IEEE Access*, vol. 10, 60622–60632, 2022.
[15] Z. Yang, X. Wang, J. Li "EIoU: An Improved Vehicle Detection Algorithm Based on VehicleNet Neural Network," *Journal of Physics: Conference Series*, vol. 1924, 1, 012001, 2021.
[16] N. Anithadevi, J. Abinisha, V. Akalya, V. Haripriya "An Improved SSD Object Detection Algorithm For Safe Social Distancing and Face Mask Detection In Public Areas Through Intelligent Video Analytics," *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1–7, 2021.
[17] Y. Xing, L. Zhong, X. Zhong "An encoder-decoder network based FCN architecture for semantic segmentation," *Wireless Communications and Mobile Computing*, vol. 2020, 2020.
[18] ALHP. Shaik, MK. Manoharan, AK. Pani, RR. Avala, C. Chen "Gaussian Mutation–Spider Monkey Optimization (GM-SMO) Model for Remote Sensing Scene Classification," *Remote Sensing*, vol. 14, 24, 6279, 2022.
[19] S. Woo, J. Park, JY. Lee, IS. Kweon "Cbam: convolutional block attention module. In proceedings of the European conference on computer vision (ECCV): 3-19," 2018.