

# Small Target Detection of Insulator Defects Based on Improved YOLOv5

Jun-Peng Wu\*, Jia-Jun Zeng

School of Electrical Engineering  
Northeast Electric Power University, Jilin 132012, China  
junpengwu80@163.com, 15695258967@163.com

\*Corresponding author: Jun-Peng Wu

Received May 24, 2023, revised October 13, 2023, accepted March 9, 2024.

---

**ABSTRACT.** *As the increasing focus on primary functions of insulators in electrical insulation and mechanical fixation in the power system, the highly accurate detection on insulator defects is of significance in maintaining the safety of power system. In this paper, an insulator defect identification method based on the enhanced You Only Look Once (YOLO) v5 model is studied, aiming at promoting the accuracy on insulator defect detections in contrast to current algorithms and solving the problems of the relative high misdetection rate. First of all, for the inaccuracy of target detection caused by small proportions of the target defective parts in the whole insulator image, multi-branch convolutions (MB-Convs) are designed as the backbone network of YOLOv5 to promote the extracting ability of the network on the feature information of small targets. Secondly, to solve the problem of small target feature information loss caused by sampling on feature pyramid network (FPN), Content-aware reassembly of features (CARAFE) -FPN structure is used to replace the standard FPN structure in YOLOv5. Finally, in order to promote the balance of positive and negative samples in the network for a further reduction in the network misdetection of insulator defects, the SoftCIoU-NMS algorithm is applied to select candidate frames in the post-processing step. According to experimental results, an average accuracy (mAP) value of insulator defect identification by the improved YOLOv5 algorithm reaches up to 81.6%, indicating the proposed detection in this paper is promoted effectively by 5.5% higher than the original.*

**Keywords:** YOLOv5, insulator defect, object detection, MB-Convs, deep learning

---

**1. Introduction.** As a crucial component in the operation of transmission lines in the power grid, insulators realize the functions of electrical insulation and equipment fixations [1]. Due to the geographical complexity of installed locations and the long-term exposure to the outdoor surroundings, the performance of insulators tends to decline or be damaged by a wide variety of adverse conditions or the unpredicted extreme weather [2], leading to the potential existence of risks and dangers and affecting the safety operation of power system. Therefore, the study on the detection of insulator defects is of great significance [3]. To date, the traditional methods have been manual inspections, which are time-consuming and have low detection accuracy and efficiency. With the technological furtherance in image processing and UAV, the methods of machine vision, including directional gradient histogram (HOG) [4] and local binary mode (LBP) [5], are increasingly applied for the advantages in high detection accuracy and fast speed in insulator defect detection, in comparison with traditional manual inspections. Chen et al. [6] devised an improved protocol ensuring that UAVs can securely transmit high-definition

data. However, these methods in practical applications are not very effective due to a range of disturbances caused by background noises in insulator defect images.

In recent years, an increasing number of domestic and foreign scholars have gradually focused on studies on insulator defect detection. Wu et al. [7] proposed an improved scheme based on PAulth to ensure the secure transmission of information over the smart grid. Zhou et al. [8] designed a defect detection system for infrared images as a solution to the inaccuracy and inefficiency of manual detection. To reduce the influence of defective insulators in the power grid, He et al. [9] applied low-frequency harmonics to position accurately the defective insulators after analyzing the relationship between the low harmonic component and insulator defect characteristics. Reddy and Mohanta. [10] adopted an adaptive vector machine to estimate the state of insulators through the extracted features from discrete orthogonal transformation (DOST). Although some prior studies have discussed the methods of identifying defective insulators, the accuracy is still relatively low. In a more recent study, Chen et al. [11,12] proposed a novel detection framework for data anomalies, Liu et al. [13] designed a pulsed infrared thermal imaging method to locate defects by comparing the temperature difference between defective and non-defective parts. Wang et al. [14] proposed an improved method of ResNeSt and area suggestion network (RPN) given the low accuracy and longtime of insulator defect detection and used a multi-scale area suggestion network to strengthen the detection of minor defects. Zhao et al. [15] extracted insulator characteristic information by improved convolutional neural networks and verified the robustness of the obtained characteristic information by SVM. Although the interference of complex background on infrared images was reduced, the detection was more time-consuming. RCNN [16], YOLO [17], and Single Shot Multi-Box Detector (SSD) [18] were improved gradually in the following studies. Miao et al. [19] developed a two-stage fine-tuning SSD algorithm, which can recognize insulators well under complex backgrounds. Ling et al. [20] first used Faster-RCNN to locate the defective parts for self-exploding glass insulators and then used U-net to segment the defective parts. Zhao et al. [21] proposed to preprocess the target area first and send it to the improved Faster-RCNN network for training because of the complex background and less feature information contained in the defect images of electrical equipment. For many small and medium-sized targets in power inspection images, Chen et al. [22] proposed an improved YOLOv2 model to increase the model's identification of small targets. Chen et al. [23] designed a YOLOv3 network [24] with super-resolution reconstruction for the insulator images were unclear, and there were few images in the data set. First, the fuzzy insulator images were reconstructed with a super-resolution to realize the amplification of the data set and then sent to the YOLOv3 network for training. To solve the low positioning accuracy on defective insulators, Ma et al. [25] replaced the original loss function in YOLOv4 with a GIOU loss function network, which accelerated the convergence speed of the network, and the accuracy was more improved to the traditional YOLOv4 by 7.37%. To meet the speed demand on real-time detection of defective insulators, Han et al. [26] reconstructed the backbone network of YOLOv4 and depth separable convolution by using lightweight GhostNet to reduce the overall number of parameters in the model. For the problem of small-sized detection targets and the complex backgrounds of insulator defects, Huang et al. [27] reduced the interference of load background on insulator defect detection in a hierarchical detection method, which used firstly the traditional YOLOv5 network to locate insulator targets and then the DenseNet201 network to further detect insulator defects. Zhu et al. [28] combined the Transformer module with YOLOv5 and added a small target detection header, significantly affecting the VisDrone2021 dataset. In more recent studies, Ahmad et al. [29] proposed an anchor frame optimization algorithm to reduce unnecessary regional suggestions for further promoting the detection accuracy

in this model. Li et al. [30] proposed an image enhancement method for illumination correction and compensation to solve the problems of uneven and inconsistent illuminations, low contrast, and insufficient specific details in outdoor insulator images and then tested insulator defects in real-time in the YOLOv5 model. Fang and Shi [31] proposed the feature information fusion method to keep the precision of small targets when lack of sufficient feature information on small targets in convolutional neural networks. Wang et al. [32] proposed a road target recognition algorithm toward smart city applications, using a migration reinforcement learning approach to improve the accuracy and real-time performance of target recognition. Figure 1 shows insulator defect detection development.

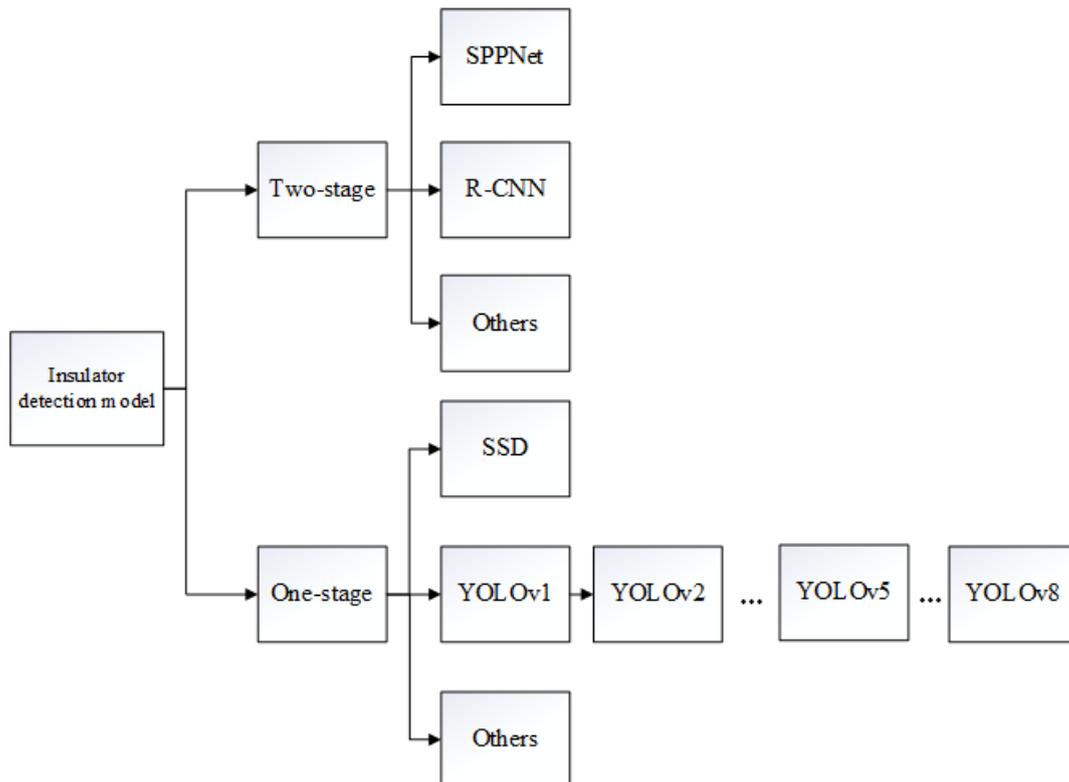


Figure 1. Insulator defect detection development

As shown in Figure 2, insulators and their defective parts are given. Compared with the traditional dataset [33], the defective parts of insulators take up a small proportion of the image. Deep convolutional neural networks retain little feature information of small targets after multiple convolutions, which is also a serious problem in insulator defect detection. In this paper, improving the total receptive field of the model helps to reduce the loss of feature information in multiple convolution operations, and the MB-Convs are designed as the backbone. In order to improve the ability of the network to extract features, the CARAFE module is used to replace the upsampling in the feature pyramid so that the upsampling increases the receptive field of the network and improves its ability to extract features. Additionally, to minimize the impact of the imbalance between positive and negative samples on detection accuracy, this paper designed the non-maximum suppression algorithm (NMS) post-processing algorithm to improve the detection accuracy.

Figure 3 displays the YOLOv5 model with improvements. Overall, the main contributions of this paper are in three areas:



Figure 2. Insulator defect examples

(1) In deep convolution, the feature extraction capability of the network for small defect targets becomes worse. In this paper, a multi-branch backbone network is designed to replace Conv in the deep backbone network to improve the network's ability to detect defective targets.

(2) The conventional upsampling method results in a significant loss of feature information for small targets. Therefore, a CARAFE-FPN feature fusion module is designed based on the CARAFE module.

(3) When the traditional NMS algorithm deals with dense small targets such as defective insulators, real targets will also be eliminated, resulting in an imbalance of positive and negative samples and a decrease in detection model accuracy. The SoftCIoU-NMS algorithm is designed to screen the prediction box, to retain more real defective insulator targets.

**2. Improvements Based on YOLOv5.** YOLOv5, as a classic single-stage algorithm, is an improved single-stage algorithm based on YOLOv3 and YOLOv4. Now it has been developed into YOLOv8, but as a new framework, YOLOv8 is not very stable. In this paper, considering the stability, real-time performance, and accuracy of the algorithm framework, YOLOv5 is adopted as the basic framework for insulator defect detection. The improved YOLOv5 mainly involves the Backbone, Neck, and Head. Backbone is mainly composed of Conv, MB-Convs, Bottleneck with 3 convolutions (C3), and Spatial Pyramid Pooling-Fast (SPPF) [34]. Conv, C3 and SPPF network structure are used, respectively to extract image feature information, reduce the computation and memory, and integrate the feature information. In the Neck, The YOLOv5 model incorporates the FPN+PAN architecture, which leverages the top-down approach of FPN to propagate deep semantic features to shallow layers [35], thereby improving semantic representation across multiple scales. Additionally, PAN facilitates the transmission of location information from shallow to deep layers [36], enhancing location ability across multiple scales.

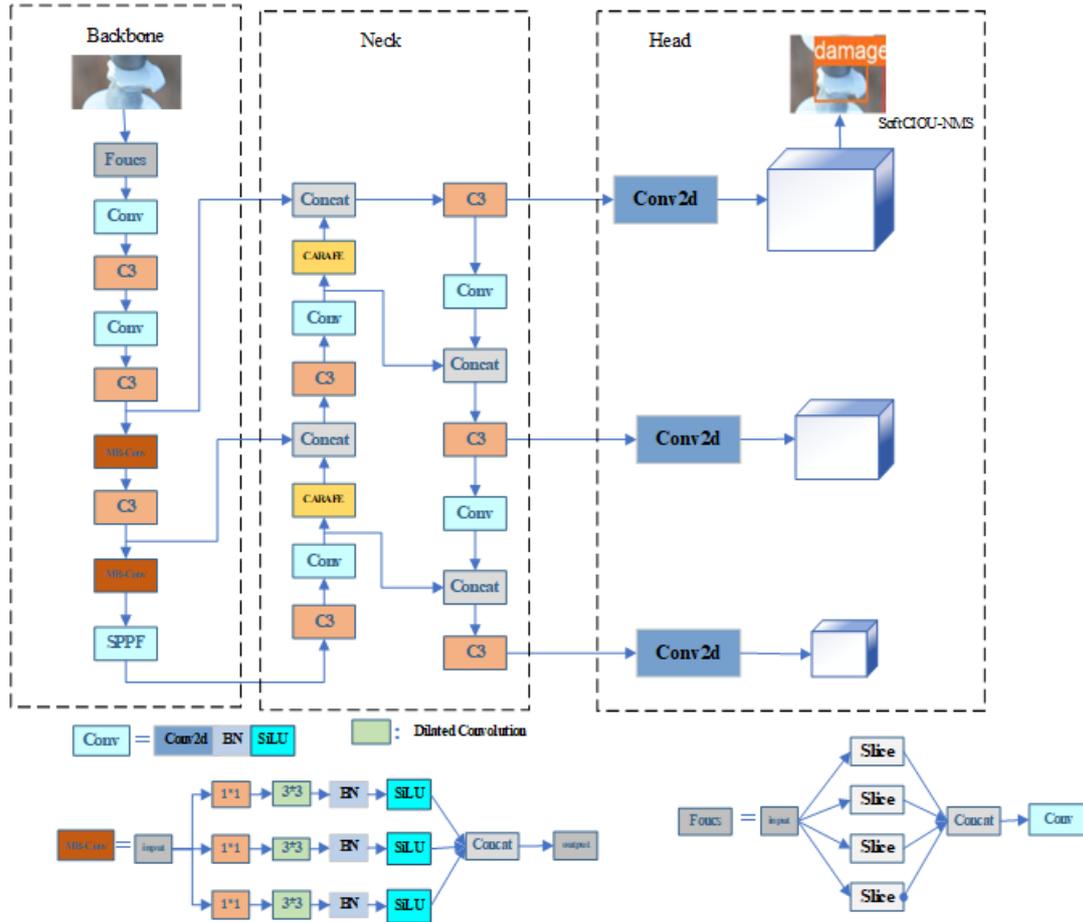


Figure 3. Improved YOLOv5 model structure

The FPN+PAN structure effectively enhances the multi-scale feature extraction capability of the network. However, the conventional upsampling method results in a significant loss of feature information for small targets. In the Head,  $80 \times 80$ ,  $40 \times 40$  and  $20 \times 20$  feature maps are output, corresponding to the detection of small-scale, medium-scale, and large-scale targets respectively. These three prediction structures are preset in all networks, and their location information and confidence information are recorded. Then the NMS is applied to screen the repeated prediction frames, eliminate the ones with low confidence and retain the ones with high confidence, and finally complete the task of target detection.

**2.1. Design of Multi-branch Backbone Network.** One of the fundamental problems with the low accuracy of insulator defect detection is that the insulator and its defective portion only account for a very small portion of the image, especially the feature information that can be extracted is also relatively small. Moreover, compared with the medium and large-scale targets, the small-scale targets have the problem of low resolution, and the features that can be extracted after multiple convolution pooling in the deep network become very few. To solve this problem, inspired by the structure of RFBNet and InceptionV2, the large receptive field can better extract features of small targets. The InceptionV2 structure ended up with multiple scales of feature information by connecting multiple branches of convolution operations, and RFBNet used a multi-scale dilated convolution to increase the range of the receptive field. Therefore, this paper designed MB-Convs as the backbone extraction network of YOLOv5 to better extract characteristic

information of insulator defects. Its basic structure is shown in Figure 1. The structure presents three branches. In each branch, a convolution operation will be carried out first to unify the number of channels of the three branches. Then, the feature information of different receptive fields will be obtained through the empty convolution layer with cavity convolution expansion rates of 1, 2, and 2, respectively. A multi-branch vacuous convolution structure is used to change the receptive field of the convolutional layer. Compared with ordinary convolution, dilated convolution can make convolution operation with any pixel with a fixed interval during operation, which brings the advantage is that a larger receptive field can be obtained under the same number of coherent convolutions. The convolution check relation between the dilated and conventional convolutions is shown in Formula (1):

$$K = k + (k - 1)(r - 1) \quad (1)$$

The actual convolution kernel size of the dilation convolution is represented by  $K$  [37],  $k$  represents the conventional convolution kernel size, and  $r$  represents the expansion rate of the dilated convolution. It can be seen from the formula that empty convolution can improve the receptive field of the obtained feature graph without increasing the parameters and the amount of computation.

**2.2. CARAFE-FPN.** Feature upsampling plays an important role in the feature pyramid. The most common upsampling strategies include bilinear interpolation, nearest neighbor interpolation, and trilinear interpolation; however, these methods only consider sub-pixel neighborhoods with a small perceptual region. In the task of insulator defect identification, insulator defect targets are dense, and these upsampling methods cannot obtain richer semantic information. Deconvolution is another approach. It uses convolution to extend the dimension of the feature map, but deconvolution uses the same convolution kernel over the entire feature map. This limits its ability to sense local feature changes and increases the amount of computation. Wang et al. proposed the lightweight upsampling operator CARAFE, which has the following advantages: including a large receptive field for better use of surrounding information, the semantic information association between kernels with upsampling based on input content and feature graphs, and not to introduce a number of parameters and computation.

As shown in Figure 4, the module consists of two main modules, namely the upsampling kernel prediction module and the feature recombination module [38]. In the given feature graph  $H \times W \times C$ . The first step of the upper sample prediction module is to predict the upper sample kernel, and then the upsampling is accomplished by the feature reorganization module to obtain a feature map that represents the upsampling rate. In the upsampling module, the feature graph is firstly compressed. And then the feature graph is expanded in dimension by pixel-shuffle to predict the feature graph with the size of  $\alpha H \times \alpha W \times k_{up}^2$  which represents the size of the upsampling kernel, and each feature value corresponds to an upsampling kernel. These upsampling cores are completely from the input feature graph and are different. Each sample kernel is mapped to the input feature map, to be obtained as a region centered on it, and then the output is derived by predicting the sample checkpoints based on this point. Multiple channels share an upper sampling kernel at the same location. In the CARAFE module, after the feature map of the upsampling kernel is predicted, an upsampling region is found on the input feature map through the mapping relationship to ensure that the location of the region corresponds to the location of the predicted up-sampled kernel one by one. In the edge position of the feature map, the number zero will be used to fill for imperfect information. In doing so, the method of reflection extension is adopted instead to retain more details of small targets and further improve the information extraction ability of small target

features. With the improved ability of the CARAFE module to collect insulator defective target details during upsampling, this paper uses the improved CARAFE module to replace the FPN structure in YOLOv5, which improves the ability of the network to identify defective insulators by replacing the CARAFE-FPN structure.

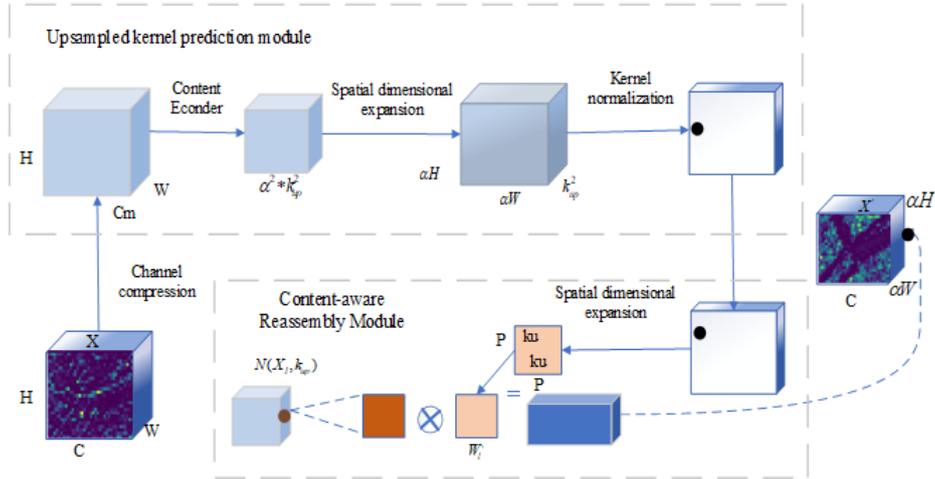


Figure 4. CARAFE module structure

**2.3. Improved Non-maximum Suppression Algorithm.** Non-Maximum suppression [39] is a necessary post-processing procedure in the target detection algorithm to eliminate redundant prediction boxes on the same object. The fundamental concept of the algorithm is as follows: First, all the boundary boxes predicted by the network are sorted from the highest score to the lowest score, and then the prediction box with the highest score is selected as the target. Intersection over Union(IOU) is used to calculate the overlap between the target and the remaining prediction boxes. Assuming that the calculated IOU value is greater than a predetermined threshold. The prediction box and target are considered to be responsible for predicting the same object at the same time, so that the boundary box will be deleted. Otherwise, it will be retained, then select the prediction box with the highest score from the undeleted prediction box as the new target, and repeat the process until all boundary boxes are identified.

Since the insulator defect targets are small-sized in high density, the traditional NMS algorithm misses the true targets with low confidence, resulting in a degradation of model accuracy. This paper uses the SoftCIoU-NMS algorithm to improve the traditional NMS algorithm. The basic idea is the same as the NMS algorithm. The basic formula of SoftCIoU-NMS is as follows:  $s_i$  indicates the score of the current detection box.  $N_1$  is the threshold of CIoU;  $n$  indicates the detection box with the highest score.  $\sigma$  is the weight coefficient. It can be seen from Formula (2) that SoftCIoU is used in this paper to replace IOU as the evaluation index of boundary frame confidence. Based on the original IOU, SoftCIoU considers the distance between the center point of the real box and the predicted box and the diagonal distance between the minimum wrapping box of the two boxes. In the SoftCIoU-NMS algorithm, the CIoU of the current and highest-score detection box is smaller than the set threshold. Output the score of the current detection box. When the CIoU of the current detection box and the highest score detection box are greater than the set threshold, the score of the current detection is multiplied by a weight function rather than being eliminated directly. This function will decay the scores of additional detection boxes that overlap with the detection box  $n$

with the highest score. When doing non-maximum suppression, the overlap of boundary frames is taken into consideration.

$$s_i = \begin{cases} s_i & C_{iou}(n, a_i) < N_1 \\ s_i \times e^{-\frac{C_{iou}(n, a_i)^2}{\sigma}} & C_{iou}(n, a_i) \geq N_1 \end{cases} \quad (2)$$

### 3. Experimental Process.

**3.1. Dataset Preprocessing.** Dexter [40] provided the defective insulator image dataset (IDID) used in this paper. There are 400 different images in the dataset. Then, 1600 images of defective insulators are obtained by a series of data magnification methods such as random flipping. The open-source data annotation tool LabelImg is used to annotate the images and generate the annotated images in VOC2017 format. The ratio of the training set, verification set, and test set is 8:1:1. Also, like YOLOv4, the Mosaic-4 data enhancement strategy is adopted by default at the input end of YOLOv5. Its basic idea is that four images are randomly scaled, randomly clipped, and randomly arranged on one image as training samples. In this paper, the Mosaic-9 data enhancement strategy is adopted to replace Mosaic-4. Nine images are randomly clipped, scaled, and mixed into one for training. Its advantage is that it dramatically increases the data information and the number of small target objects. In the normalization operation, nine images are calculated without relying on batch processing parameters, which reduces the calculation amount and thus improves the detection efficiency of the model.

**3.2. Experimental Environment.** The experimental operating system used in this study is Windows 11, Python 3.8 environment, and Pytorch framework. GeForce RTX 3060 Laptop GPU with 6G video memory and 16G memory is the type of graphics card used. The training process uses a 640x640 pixel image with 16 batch training data points, a training momentum of 0.937, a weight attenuation of 0.0005, and an initial learning rate of 0.01.

**3.3. Performance Evaluation.** Precision ( $P$ ), Recall ( $R$ ), False detection ( $F$ ), Average Precision ( $AP$ ), and Average Precision Mean ( $mAP$ ) are used to measure the performance of the improved model. The full calculation is made available as follows:

$$P = \frac{TP}{TP + FP} \quad (3)$$

$$R = \frac{TP}{TP + FN} \quad (4)$$

$$AP = \int_0^1 p(R) dR \quad (5)$$

$$mAP = \frac{\sum_{i=0}^n AP(i)}{n} \quad (6)$$

$$F_1 = \frac{2 * P * R}{P + R} \quad (7)$$

Where True positives ( $TP$ ) is the number of positive samples accurately recognized by the model as positive samples, False positives ( $FP$ ) is the number of negative samples incorrectly recognized as positive samples, True negatives ( $TN$ ) is the number of negative samples accurately recognized by the model as negative samples, False negatives ( $FN$ ) is the number of positive samples incorrectly identified by the model as negative samples and

$n$  is the total number of categories of detection targets;  $AP$  is the area under the Precision-Recall curve;  $mAP$  is the average value of detected  $AP$ . The increasing number of  $mAP$  presents closely relates to the excellence of network performance; as  $mAP$  increases, the network performance is promoted. Since the two indexes of precision rate  $P$  and recall rate  $R$  are contradictory, it is impossible to achieve a double high. The  $F_1$  score is balanced according to their balance point, considering both precision and recall.

**4. Experimental Results and Analysis.** In this paper, based on the original YOLOv5 model, MB-Convs are used as the backbone of the network to realize the precise location of insulator faulty components, and the enhanced CARAFE module replaces the FPN structure in YOLOv5. The post-processing technique makes use of the SoftCIoU-NMS algorithm to enhance the capacity to identify damaged insulators.

Finally, the experimental comparison results are shown in Table 1. It can be seen from the table that both the original model and the improved model have strong positioning ability for the insulator itself, but the original YOLOv5 model has strong positioning ability for the insulator defects as follows: The positioning ability of flashover and damage is poor. Compared with the original YOLOv5, the accuracy of improved YOLOv5 in flashover increases by 8.7%, while the accuracy  $P$  and recall rate  $R$  rises by 6.2% and 13.8%, respectively. Accuracy rises by 6% in damage, while accuracy  $P$  and recall  $R$  rise by 2.9% and 5.2%, respectively. Last but not least, the recall rate of  $R$  grows by 6.4%, the accuracy of  $P$  increases by 4.8%, and the average accuracy of  $mAP$  increases by 5.5%. The results demonstrate that the improved YOLOv5 method in this study has a better detection for the small targets in insulator flaws.

Table 1. The improved model is compared with the original model

Category	Original YOLOv5			Improved YOLOv5		
	mAP@0.5(%)	P(%)	R(%)	mAP@0.5(%)	P(%)	R(%)
Insulator	97.5	95.8	94.2	97.6	96.3	94.1
Flashover	61.6	70.2	66.3	70.3	76.4	80.1
Damage	70.9	76.3	72.1	76.9	79.2	77.3
ALL	76.1	78.7	79.6	81.6	82.5	86.2

Figure 5 to Figure 6 respectively show both  $F_1$  value and  $P-R$  curve about the original and improved models. The figure shows that the modified model has improved the  $F_1$  value and  $P-R$  curve to the varying degrees when compared with the original model.

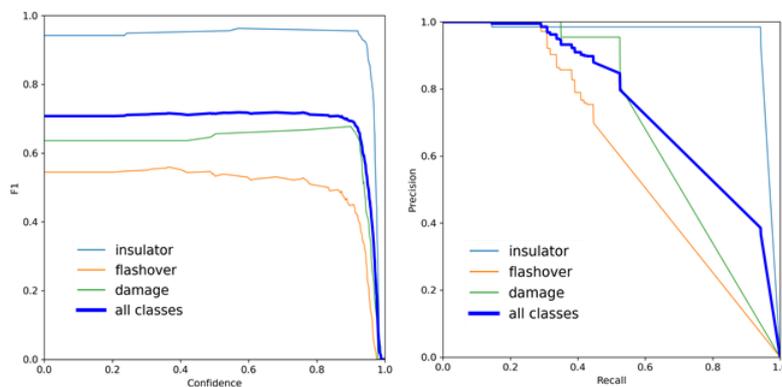


Figure 5.  $F_1$  curve and  $P-R$  curve of the original YOLOv5

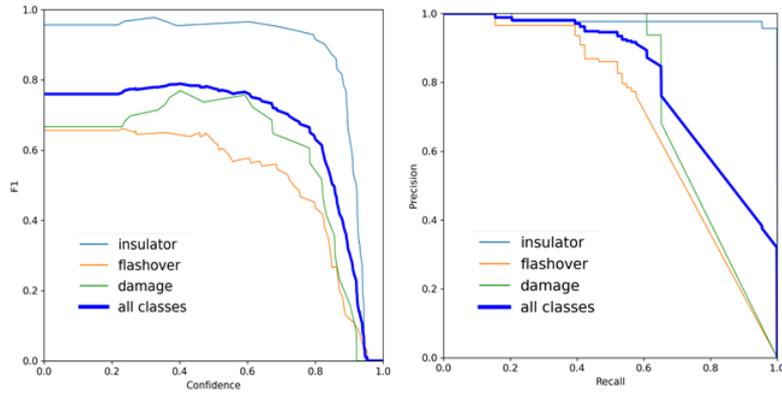


Figure 6.  $F_1$  curve and  $P - R$  curve of the improved YOLOv5

Figure 7 depicts the Grad-CAM diagrams of the original and improved YOLOv5 models during the testing phase. Among them, (a) is the original image, (b) is the insulator defect detection Grad-CAM diagrams of the original YOLOv5, and (c) the insulator defect detection Grad-CAM diagrams of the improved YOLOv5. The varying hues in the image represent different gradients between the current layer and the output layer, with the redder portion representing the focus of the network. As seen in the image, the color of the improved YOLOv5 model is darker in the minor defect target area when compared with the original model. These locations are given more importance, and the model is more likely to discover the damaged parts. The findings demonstrate that the revised model can detect insulator defects more accurately and reduce missing and erroneous detection.

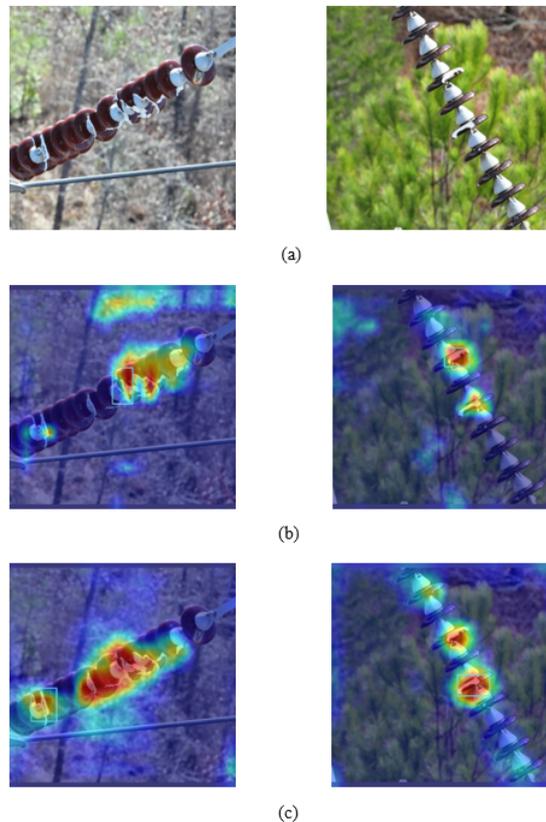


Figure 7. Heatmap of the experimental results

Figure 8 shows the insulator defect detection results of the original YOLOv5 and the improved YOLOv5. Among them, (a) is the insulator defect detection effect drawing of the original YOLOv5, and (b) is the insulator defect detection effect drawing of the improved YOLOv5. The original YOLOv5 method has missed detection and false detection because of the small number of insulator defect targets in the image, some of which are highly dense, and the background is relatively complicated. However, in this paper, MB-Convs as the backbone network, CARAFE upsampling operator and SoftCIOU-NMS are used in the algorithm to enhance the ability of the network to learn the feature information of small targets. The improved model significantly reduces the number of missing insulator defect parts. The method in this paper can accurately distinguish the target and show the advantages of this method in the provided results.

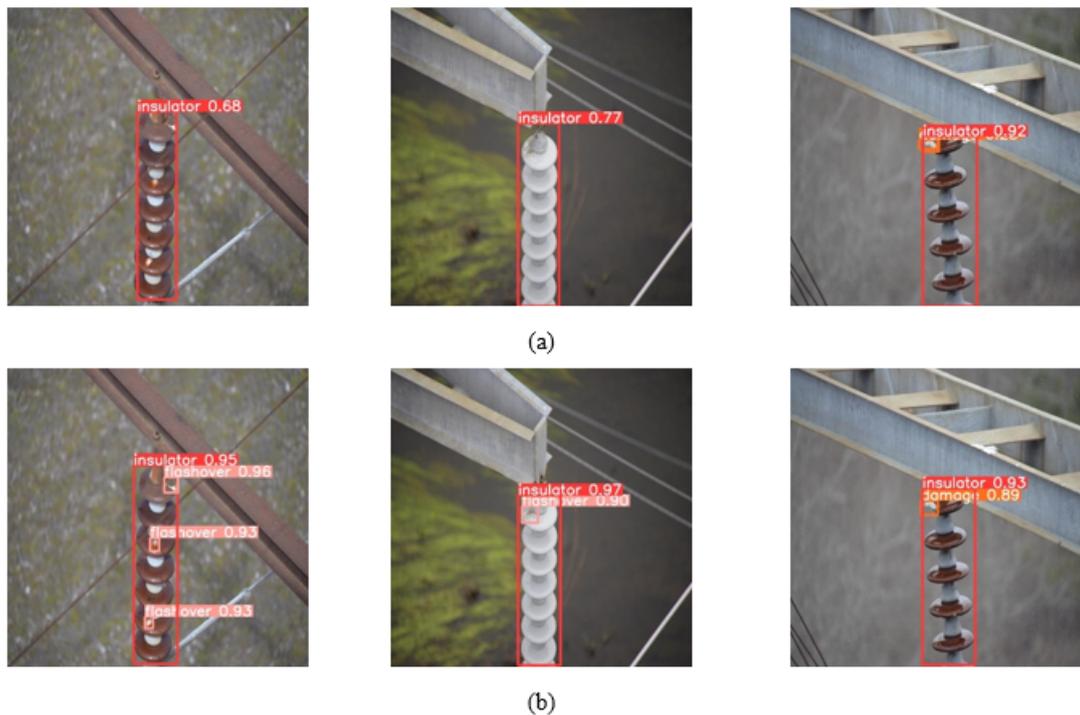


Figure 8. Insulator defects detection results

The ablation experiment is carried out in this paper to verify the performance of the improved YOLOv5 model proposed in insulator defect detection. The improved mechanisms are added to the original structure and trained to verify whether each improved mechanism is effective for the model. Table 2 displays the results.

The table shows the original YOLOv5s detection precision. YOLOv5s  $mAP$  increases by 5% when MB-Convs are used as the backbone network because MB-Convs improve the ability to extract small defect targets from deep convolution. When the CARAFE upsampling operator is used to replace the nearest interpolation in the original network, the  $mAP$  value is improved by 4.1%. When SoftCIOU-NMS is used as the post-processing algorithm, the  $mAP$  value of the network is improved by 2%. Two different improvement strategies are added to the original YOLOv5 network, and the  $mAP$  value is increased by 3.7%, 5.1% and 3.1%, respectively; experiments show that multi-branch backbone, SoftCIOU-NMS, and CARAFE sampling operator network have higher precision. Finally, to multi-branch backbone network fusion, SoftCIOU-NMS, and sampling on the CARAFE operator, after getting the final accuracy, compared with the original YOLOv5 model, the  $mAP$  value increases by 5.8%. The algorithm in this paper synthesizes the advantages

Table 2. Comparison table of ablation experiments

Number	MB-Convs	SoftCIOU-NMS	CARAFE	mAP@0.5(%)
1				76.1
2	✓			81.1
3		✓		78.1
4			✓	80.2
5	✓	✓		79.8
6		✓	✓	81.2
7	✓		✓	79.4
8	✓	✓	✓	81.6

of the above modules,  $mAP$  value can reach 81.9% to achieve accurate positioning of defective insulators.

Non-maximum suppression algorithms have varying effects on the detection accuracy of defective insulators; different non-maximum suppression methods are applied on the basis of the original YOLOv5 model, after which the detection is performed correspondingly. Table 3 displays the experimental results, which show that the technique used in this research has the best effect on identifying defective insulators. The average recognition accuracy has increased by 1.3% and 0.9%, respectively, when compared Soft-NMS with CIUO-NMS. This is due to the fact that SoftCIUO-NMS considers the overlap area, center distance, and aspect ratio in addition to immediately eliminating candidate boxes that are larger than the IOU threshold. In insulator defect detection, the problem of an imbalance between positive and negative samples is resolved, and the model's detection precision increases.

Table 3. Comparison between algorithms of different NMS

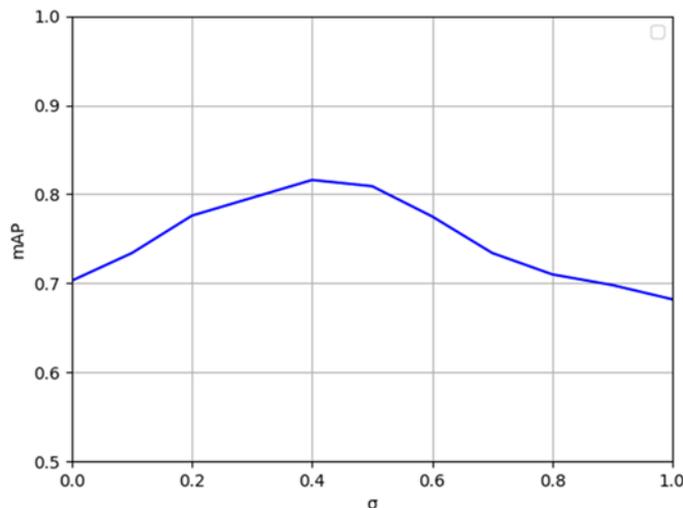
Method	$mAP@0.5$ (%)
Soft-NMS	76.8
CIUO-NMS	77.2
DIOU-NMS	76.4
SoftCIUO-NMS	78.1

As the number of MB-Convs have different impacts on model accuracy, a control experiment is designed in this paper. In the original YOLOv5 model, ordinary volumes in the backbone network are replaced by different numbers of MB-Convs, and the detection accuracy of the model is obtained, as shown in Table 4. It can be seen from the table that the performance of the model is related to the number of MB-Convs. When the last two common convolutions in the backbone network replace the two MB-Convs, the average accuracy of the model reaches the highest 81.6%, and, with the increase of the number of substitutions, the complexity of the model and the amount of floating-point computation also increase. When all the convolutions in the backbone network are replaced by the MB-Convs, the volume of the model reaches 8.5M.

In SoftCIUO-NMS, different weight coefficients have different effects on the location accuracy of insulator defects. Based on the enhanced YOLOv5, the improved YOLOv5 model is drawn with varied weight coefficients to show the detection accuracy of the defective insulators. Figure 9 shows that when the parameter is 0.4, the model achieves the best accuracy of 81.6%. If this value is too high or too low, the detection accuracy of the upgraded model decreases. As a result, 0.4 is chosen as the weight coefficient in this paper.

Table 4. Multiple branch convolutions with different numbers

Number	$mAP@0.5$ (%)	GFLOPs	Parameter/M
1	80.2	16.9	7.5
2	81.6	17.9	7.8
3	79.8	18.8	8.2
4	77.5	19.8	8.5

Figure 9. The accuracy of the improved model under different  $\sigma$ 

A horizontal comparison is conducted between the improved model and the most popular target detection models currently in use, including Faster-RCNN, YOLOv3, SSD, and YOLOv4, to confirm the improved model's effect. In this paper, Faster-RCNN, YOLOv3, SSD, and YOLOv4 are trained with the same hyper-parameters for comparison by using the same dataset. According to Table 5, the improved YOLOv5 algorithm enhances model accuracy by 1.4% when to the traditional two-stage approach Faster-RCNN. The improved model volume and floating-point calculation amount are also significantly lower than those of Faster-RCNN. The SSD network has a complex model structure and the lowest detection accuracy, only 74.4%. Despite about one-fifth of the parameters of the YOLOv3 and YOLOv4 models, the accuracy is increased by 5.7% and 4.4%, respectively. This paper also compares the more popular detection model, TPH-YOLOv5, and from the results, it can be seen that the detection model proposed in this paper is more effective in the detection of insulator defects and less intensive computationally.

Table 5. Comparison of different methods

Method	$mAP@0.5$ (%)	Parameter/M	GFLOPs	FPS
Faster-RCNN	80.2	72.4	27.1	9.8
SSD	74.4	25.8	33.4	18.3
YOLOv3	75.9	62.5	156.4	27.2
YOLOv4	77.2	52.6	119.7	21.3
Improved YOLOv5	81.6	10.3	19.4	26.6
TPH-YOLOv5 [28]	80.4	19.8	212.6	22.4

**5. Conclusion.** Due to the small proportion of defective parts in the image, an improved YOLOv5 model is designed in this paper to solve the low accuracy of the model in identifying insulator defects. Firstly, because of the poor feature information extraction ability of the original network for insulator defects, MB-Convs are designed to replace the common convolution in the YOLOv5 backbone network. Secondly, since the features of small targets tend to be stronger in shallow feature maps and weaker or even disappear in deep feature maps, CARAFE-FPN is used instead of the original FPN structure for feature fusion so as to reduce the loss of feature information. Finally, SoftCIoU-NMS is designed in the post-processing algorithm for defect-intensive problems. After training, the final model is obtained. Compared with other mainstream detection models, the accuracy has been increased to different degrees. Therefore, the improved model has obvious advantages in the identification and location of insulator defects. However, the improved model cannot meet the real-time detection of defective insulators, and the real-time detection of defective insulators will be the next focus of attention.

**Acknowledgements.** The research is supported by: Jilin Province Science and Technology Development Plan Project(20200403075SF) and Jilin Provincial Department of Education Project (JJKH20240148KJ).

## REFERENCES

- [1] H. J. Liu, S. P. Geng, J. Wang, "Deterioration analysis of porcelain insulator in UHV AC transmission line," *Insulators and Surge Arresters*, vol. 310, no. 6, pp. 159-164, 2022.
- [2] Y. H. Li, G. S. Wang, X. M. He, "Analysis of breakdown defect and preventive measures of composite pillar insulator," *Guangxi Electric Power*, vol. 44, no. 2, pp. 100-104, 2021.
- [3] Z. B. Qiu, J. J. Ruan, D. C. Huang, "Analysis and Experimental study on Aging form of suspended porcelain insulators for transmission lines," *High Voltage Engineering*, vol. 42, no. 4, pp. 1259-1267, 2016.
- [4] T. Yan, G. Yang, J. Yu, "Feature fusion based insulator detection for aerial inspection," 2017 36th Chinese Control Conference (CCC), pp. 10972-10977, 2017.
- [5] L. Yang, X. Jiang, Y. Hao, L. Li, H. Li, R. Li, B. Luo, "Recognition of natural ice types on in-service glass insulators based on texture feature descriptor," *IEEE Transactions on Dielectrics Electrical Insulation*, vol. 24, no. 1, pp. 535-542, 2017.
- [6] T.-Y. Wu, X.-l. Guo, Y.-C. Chen, S. Kumari, C.-M. Chen, "Amassing the Security: An Enhanced Authentication Protocol for Drone Communications over 5G Networks," *Drones*, vol. 6, no. 1, 10, 2022.
- [7] T.-Y. Wu, Y.-Q. Lee, C.-M. Chen, Y. Tian, N. A. Al-Nabhan, "An Enhanced Pairing-based Authentication Scheme for Smart Grid Communications," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-13, 2021.
- [8] S. P. Zhou, X. Lei, B. C. Qiao, H. Hou, "Research on Insulator Fault Diagnosis and Remote Monitoring System Based on Infrared Images," *Procedia Computer Science*, vol. 109, pp. 1194-1199, 2017.
- [9] S. Han, R. Hao, J. Lee, "Inspection of Insulators on High-Voltage Power Transmission Lines," *IEEE Transactions on Power Delivery*, vol. 24, no. 4, pp. 2319-2327, 2009.
- [10] M. J. B. Reddy, K. C. B. D. K. Mohanta, "Condition monitoring of 11 kV distribution system insulators incorporating complex imagery using combined DOST-SVM approach," *IEEE Transactions on Dielectrics Electrical Insulation*, vol. 20, no. 2, pp. 664-674, 2013.
- [11] W. S. Gan, L. L. Chen, S. C. Wan, J.H Chen, C.-M. Chen, "Anomaly Rule Detection in Sequence Data," *IEEE Transactions on Knowledge and Data Engineering*, pp. 6261-6270, 2021.
- [12] M. T. Wu, Q. Teng, S. Huda, Y.C. Chen, C.-M. Chen, "A privacy frequent itemsets mining framework for collaboration in IoT using federated learning," *ACM Transactions on Sensor Networks*, vol. 19, no. 2, 27, 2023.
- [13] L. Liu, H. Mei, L. Wang, C. Zhao, Z. Guan, "Pulsed infrared thermography to inspect the internal defects of composite insulators," 2017 IEEE Electrical Insulation Conference (EIC), pp. 476-470, 2017.

- [14] S. Wang, Y. Liu, Y. Qing, C. Wang, T. Lan, R. Yao, "Detection of Insulator Defects With Improved ResNeSt and Region Proposal Network," *IEEE Access*, vol. 8, pp. 184841-184850, 2020.
- [15] Z. Zhao, X. Fan, G. Xu, L. Zhang, Y. Qi, K. Zhang, "Aggregating Deep Convolutional Feature Maps for Insulator Detection in Infrared Images," *IEEE Access*, vol. 5, pp. 21831-21839, 2017.
- [16] S. Q. Ren, K. M. He, R. B. Girshick, J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 39, pp. 1137-1149, 2015.
- [17] J. Redmon, A. Farhadi, "YOLO9000: Better, Faster, Stronger," *IEEE Conference on Computer Vision & Pattern Recognition*. IEEE, pp. 6517-6525, 2017.
- [18] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, E. Reed, C. Y. Fu, A. C. Berg, "SSD: Single Shot MultiBox Detector," *European Conference on Computer Vision*, pp. 21-27, 2015.
- [19] X. Miao, X. Liu, J. Chen, S. Zhuang, J. W. Fan, H. S. Jiang, "Insulator Detection in Aerial Images for Transmission Line Inspection Using Single Shot Multibox Detector," *IEEE Access*, pp. 1-1, 2019.
- [20] Z. Ling, R. C. Qiu, Z. Jin, Y. Zhang, C. Lei, "An accurate and real-time method of self-blast glass insulator location based on faster R-CNN and U-net with aerial images," *CSEE Journal of Power and Energy Systems*, vol. 5, no. 4, pp. 474-482, 2019.
- [21] W. Q. Zhao, M. F. Xu, X. F. Cheng, Z. B. Zhao, "An Insulator in Transmission Lines Recognition and Fault Detection Model Based on Improved Faster RCNN," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-8, 2021.
- [22] Q. Chen, Y. Gao, Y. Peng, J. Zhang, K. Sun, "Accurate Object Recognition for Unmanned Aerial Vehicle Electric Power Inspection using an Improved YOLOv2 Algorithm," *2019 IEEE Fourth International Conference on Data Science in Cyberspace (DSC)*, pp. 610-617, 2019.
- [23] H. Chen, Z. He, B. Shi, Zhong. T, "Research on Recognition Method of Electrical Components Based on YOLOV3," *IEEE Access*, vol. 7, pp. 157818-157829, 2019.
- [24] J. Redmon, A. Farhadi, "YOLOv3: An Incremental Improvement," *arXiv e-prints*, vol. 1804.02767, 2018.
- [25] B. Ma, Y. Fu, C. Wang, J. Li, Y. Wang, "A high-performance insulators location scheme based on YOLOv4 deep learning network with GDIoU loss function," *IET Image Processing*, vol. 16, no. 4, pp. 1124-1134, 2022.
- [26] G. J. Han, L. Zhao, Q. Li, S. D. Li, R. J. Wang, Q. W. Yuan, M. He, S. Q. Yang, L. Qin, "A Lightweight Algorithm for Insulator Target Detection and Defect Identification," *Sensors*, vol. 23, no. 3, 2023.
- [27] Z. Huang, S. Hu, L. Zhang, "Fault Detection of insulator in distribution network Based on YOLOv5s Neural Network," *2022 International Conference on Artificial Intelligence and Computer Information Technology (AICIT)*, pp. 1-5, 2022.
- [28] X. Zhu, S. Liu, X. Wang, Q. Zhao, "TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios," *arXiv*, pp. 2778-2788, 2021.
- [29] M. Ahmad, M. Abdullah, D. Han, "Small Object Detection in Aerial Imagery using RetinaNet with Anchor Optimization," *2020 International Conference on Electronics, Information, and Communication (ICEIC)*, pp. 1-3, 2020.
- [30] Y. Li, M. Z. Ni, Y. F. Lu, "Insulator defect detection for power grid based on light correction enhancement and YOLOv5 model," *Energy Reports*, vol. 8, pp. 807-814, 2022.
- [31] P. Fang, Y. Shi, "Small Object Detection Using Context Information Fusion in Faster R-CNN," *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*, pp. 1537-1540, 2018.
- [32] K. Wang, C.-M. Chen, M. S. Hossain, G. Muhammad, S. Kumar, S. Kumari, "Transfer reinforcement learning-based road object detection in next generation IoT domain," *Computer Networks*, vol. 193, 108078, 2021.
- [33] T. Y. Lin, M. Maire, S. Belongie, J. Hays, C. L. Zitnick, "Microsoft COCO: Common Objects in Context," *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*. Springer International Publishing, pp. 740-755, 2014.
- [34] K. He, X. Zhang, S. Ren, J. Sun, "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 37, no. 9, pp. 1904-16, 2014.
- [35] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, S. Belongie, "Feature Pyramid Networks for Object Detection," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2117-2125, 2017.

- [36] S. Liu, L. Qi, H. Qin, J. Shi, J. Jia, "Path Aggregation Network for Instance Segmentation," Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 8759-8768, 2018.
- [37] F. Yu, V. Koltun, T. Funkhouser, "Dilated Residual Networks," IEEE Computer Society, pp. 472-480, 2017.
- [38] J. Q. Wang, K. Chen, R. Xu, Z. W. Liu, C. C. Loy, D. D. Lin, "CARAFE: Content-Aware Re-Assembly of Features," IEEE/CVF International Conference on Computer Vision, pp. 3007-3016, 2019.
- [39] A. Neubeck, L. Gool, "Efficient Non-Maximum Suppression," International Conference on Pattern Recognition, vol. 3, pp. 850-855, 2006.
- [40] L. Dexter, "Insulator Defect Image Dataset - Version 1.2: Readme EPRI," Palo Alto, CA, 2020.