Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105 (Online)
Volume **7**, Number **4**, November **2022**

Fault Detection Method for Insulators Using Improved YOLOv4

Li-Quan Zhao*

Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology, Ministry of Education
Northeast Electric Power University
168 Changchun Road, Jilin, 132012, China
zhaoliquan@neepu.edu.cn

Zi-Cong Jiang

Graduate school of Information Science and Electrical Engineering
Kyushu University
Fukuoka, 819-0395, Japan
jiangzicong1234@gmail.com

Zi-Ming Teng

College of information and communication
Jilin University
Changchun, 130012, China
tengziming2002@163.com

Yan-Fei Jia

School of Electrical and Information Engineering
Beihua University
Jilin 132013, China
jiayanfei@163.com

*Corresponding author: Li-Quan Zhao

Received March 29, 2022, revised May 11, 2022, accepted July 24, 2022.

Abstract. *A defective insulator detection method based on improved You Only Look Once version 4 (YOLOv4) is proposed to improve the precision of defective insulator detection. This method designs a dual-branch attention block by simulating the attention process of the human eye, and introduces it into the backbone of YOLOv4 to extract richer information on defective insulator features. To make the output feature map of shallow layer network fuse more different scale features, a path aggregation network has been designed. To improve detection accuracy, the spatial pyramid pooling has been replaced by atrous spatial pyramid pooling. A dynamic non-maximum suppression method is proposed to improve defective insulator detection accuracy. It employs an adaptive threshold control strategy and weighted operation to full use of the predicted box. We compare our method with defective insulator detection methods based on YOLO and R-CNN. Our simulation results show that the proposed method is more accurate than defective insulator detection methods based on YOLO, and has a faster detection speed than methods based on R-CNN. Compared with the defective insulator detection method based on R-CNN that has the highest mean average precision, the mean average precision of the proposed method is lower, but its detection speed is faster. Compared with YOLOv4, its detection speed is slower, but the mean average precision is higher. The proposed method strikes a satisfactory tradeoff between detection speed and mean average precision.*
Keywords: Deep learning, Object detection, Defective insulator, Path aggregation network

1. **Introduction.** The power system consists of power generation, transformation, transmission, and distribution. Transmission lines are the infrastructure for power transmission in the power system and are the key to improving the stability and safety of the power supply. The main components that constitute transmission are conductors, lightning cables, fixtures, insulators, towers, tie wires and foundations, grounding devices, etc. The insulators are essential components that are used to support and sustain electrical transmission lines and ensure a degree of insulation between transmission lines and the ground in high-voltage power transmission systems. If the insulator fails, it will directly affect the line insulation, and then a permanent ground fault will occur [1]. Therefore, it is very important and essential to detect the insulator strings and timely deal with defective insulator string to improve the safe and stable operation of power systems. The insulator strings are exposed to wind, rain, sunlight, and other poor working conditions for a long time and are susceptible to self-exploding or going missing, which in turn will cause single-phase grounding or short-circuiting between phases [2]. This will cause power transmission interruption. To ensure safety and stability in high-voltage power transmission systems, it is necessary to detect insulator defects timorously. Manual inspection is the most commonly used transmission line inspection method. The inspector uses binoculars to observe the transmission line insulation through the human eye and to record the defective insulator conditions. However, due to the continuous expansion of the scale of power grids, more and more high-voltage transmission towers and longer transmission lines require a large number of inspectors to inspect the transmission lines. Transmission lines are generally located in remote areas, and some towers are erected in high mountains. Inspectors need to go over high mountains for inspection, resulting in low efficiency of transmission line inspection. Therefore, the traditional manual transmission line inspection methodologies can no longer satisfy this requirement.

With the deployment of unmanned aerial vehicle technology, the transmission line inspection based on the unmanned aerial vehicle has become an important means of transmission line inspection, and these are widely used by power supply companies. Transmission line inspectors use the unmanned aerial vehicle to inspect transmission lines under the mountains, and line inspections can be realized without going over the mountains. This greatly improves the efficiency of line inspection, reduce the labor intensity of inspection personnel. However, the transmission line inspection based on unmanned aerial vehicle still requires the inspector to view video obtained by unmanned aerial vehicle to detect defective insulators. To further improve the efficiency of transmission line inspection, the transmission line inspection based on unmanned aerial vehicle and deep learning is proposed. Detection method based on the deep learning method has been introduced into transmission equipment fault detection systems to supplement or replace human detection. They use the deep learning method to automatically detect transmission lines, normal insulator strings, and defective insulators without human intervention. Deep learning is one of the research directions of artificial intelligence. It has been widely used in trade strategy [3], transportation mode detection [4], image caption [5] and wind power generation forecast [6]. Although there exists a wide range of transmission equipment faults, this paper focuses only on defective insulator detection, which concerns the most common problem affecting transmission insulators.

The traditional defective insulator detection methods require an artificially designed feature extractor to extract features. The robustness of the designed feature extractor and the quality of extracted features tend to below. Advances in deep learning technology have led to the development of convolution neural networks that can be used to extract

features without relying on an artificially designed feature extractor. Many object detection methods based on convolution neural networks have been applied in the detection of defective insulators. They have better feature expression ability and higher detection accuracy than traditional detection methods. Object detection methods based on deep learning can be divided into two categories: one-stage detection method and two-stage detection method. For two-stage method, it divides the detection process into two stages. It generates region proposals in the first stage, regresses the bounding box and on candidate regions, and classifies the objects in the second stage. For the one-stage method, directly generates detection boxes and classifies the objects without generating region proposals. The one-stage method has a faster detection speed and lowers detection accuracy than the two-stage method. The transmission line inspection based on unmanned aerial vehicle requires real-time. The detection of the one-stage method is about ten times faster than two-stage. Therefore, we select the YOLOv4 method to detect the defective insulator. YOLOv4 is the latest version of the YOLO serial methods [7]. YOLOv4 is also the only later version of YOLO endorsed by Joseph Redmon, the inventor of the YOLO method [8]. Compared with the two-stage detection method, although the YOLOv4 has a faster detection speed, it also has a lower precision. The performance of the deep learning method directly affects insulator string and defective insulator detection. Therefore, we propose an improved YOLOv4 to improve the precision of defective insulators with an acceptable detection speed.

The main contributions of this paper are as follows:

1. To reduce redundant semantic information and thus reduce the influence of redundant information on defective insulator detection accuracy, a novel Dual-Branch Attention Block is designed and introduced into the backbone. This can make the backbone extract more useful feature information from complex semantic information.

2. To make the output feature map of shallow layer network fuse more different scale features, a new path aggregation network is proposed. To further improve accuracy, the ASPP (atrous spatial pyramid pooling) structure is used instead of SPP (spatial pyramid pooling) structure in the path aggregation network. ASPP can increase the respective field and reduce loss of information caused by the pooling operation.

3. To improve the accuracy of defective insulator detection, a dynamic non-maximum suppression method is proposed. This uses an adaptive threshold control strategy and weighted operation to make full use of the predicted box.

In this section, we have outlined the theoretical background of defective insulator detection and our contributions to the field. In Section 2, we review related work on defective insulator detection based on deep learning. In Section 3, we explain our proposed method in greater detail. In Section 4, we illustrate and discuss our experimental results. In Section 5, we provide a summary of our experimental work.

2. **Related work.** Object detection method based on deep learning has been widely applied in insulator and defective insulator detection. To detect the defective insulators and birds nest in high voltage lines, a deep convolution neural network on the basis of Faster-RCNN was proposed [9]. It transformed the classification problem of damage insulators and birds nest into detection and recognition. It also proposed to use the ImageNet dataset to pre-train the convolution neural network to reduce the influence of lacking many natural bird nests and damaging insulators images. Zhao et al. [10] also proposed an insulator detection method based on improved Faster-RCNN. To make the Faster-RCNN more suitable for detecting the insulators with different scales, aspect ratios, and transformation, a new region proposal network was proposed. It combined six scales with five different aspect ratios to obtain 30 anchors. Besides, it also proposed

a non-maximum suppression method to obtain more accurate bounding boxes. Lin et al. [11] proposed a two-stage self-blast glass insulator detection method. It firstly used Faster-RCNN to locate the glass insulators from the aerial images captured by UAV and then cropped the detected insulators. Secondly, U-net that is a state-of-art segmentation module, was used to judge whether the glass insulators had self-blast insulator. It had higher accuracy and a faster detection speed. Tao et al. [12] proposed cascading network that contains insulator string localizer network and insulator defect detector network, which were designed on the basis of a region proposal network that is also used in faster R-CNN. It used an insulator localizer network to detect the insulator string and cropped the detected insulator string image. The insulator defect detector network was used to judge whether the insulator string contained a defective insulator. Besides, it also constructed a public Chinese power line insulator dataset containing normal insulators and defective insulators. Liu et al. [13] proposed a Box-Point detector to detect defective insulators. The detector was composed of a deep convolutional neural network and two parallel branches: center box heads and endpoint heads. Each branch has three different heads, and each head is constructed by two convolutional layers with different kernel sizes and corresponds to a specific attribute of the defective insulator. The deep convolutional neural network was applied to extract insulators feature. The center box head was used to detect the rectangle region of defective insulator. The end point head was used to supervise the detector to lean insulators image features. Zhao et al. [14] proposed an improved Faster RCNN method by introducing new feature pyramid networks and used it to detect the normal insulator strings and defective insulators. It also firstly detected all insulator strings and secondly detected the defective insulator from the detected insulator strings.

The Faster-RCNN method can be seen as one of the R-CNN serial methods. The object detection method based on deep learning can be divided into two classifications that are one-stage detection method and two-stage detection method. The two-stage detection method solves the classification problem of the object in the first stage and searches the bounding box of the object in the second stage. The one-stage method solves classification and location problems at the same time. The two-stage method focus on the accuracy of the one-stage method focuses on detection speed. The R-CNN serial methods belong to two-stage detection methods. They have higher accuracy and slower detection speed than the one-stage method. On the contrary, the one-stage method has a faster detection speed and lower accuracy than the two-stage method. YOLO serial methods are typical one-stage detection methods. In this paper, we focus on real-time defective insulator detection, which is more important in actual transition line inspection. Therefore, we mainly introduce YOLO serial methods and their applications in defective insulator detection in the following.

YOLO method was firstly proposed by Joseph Redmon and Ali Farhadi et al. in 2016 [8]. It considers the recognition problem as a regression problem rather than a classification problem and predicts the location and category of the object at the same time by the convolutional neural network. It has a faster detection speed with acceptable accuracy. Due to its advantages, it attracts much attention from many scholars and becomes an essential branch of object detection research. In 2016, Joseph Redmon et al.proposed the YOLOv2. It used Darket19 as the backbone and many strategies, such as adding Batch Normalization in all convolution layers, using clustering to automatically generate the appropriate bounding box prior and multi-scale training, etc., to improve detection accuracy. YOLOv3 is the last version proposed by Joseph Redmon in 2018. It used Darknet53 as the backbone to extract features. Compared with Darket19, the Darknet53 increased the number of network layers and introduced cross residual network

structure of the residual network. Besides, it also used the LeakyReLU function as activate function to improve detection accuracy further. To improve the accuracy of YOLOv3, many methods were proposed. Ma et al. used the ShuffleNetv2 instead of DarkNet53 as backbone and GIOU function as loss function in 2020 [15]. It has higher detection accuracy than YOLOv3 in detecting collapsed buildings after the earthquake. Yang et al. proposed learnable semantic fusion and global context block on the basis of YOLOv3 to make full use of the output of the feature extraction network [16]. The method named GC-YOLOv3 has higher accuracy than YOLOv3 in the COCO dataset and PASCAL VOC dataset with a small amount of additional computational cost. Liu et al. designed Dense Block with multi-scale feature fusion. They used it instead of some residual network with lower resolutions in the backbone of YOLOv3 to obtain abundant semantic information of upper and lower layers [17]. It has higher accuracy than YOLOv3 in detecting transmission line insulators.

Although Joseph Redmonet al. withdrew the research on the computer version, many scholars still try to improve the YOLO serial methods. Alexey et al. [7] proposed the YOLOv4 method, that is the latest version of YOLO serial methods in 2020, which is also only admitted by Joseph Redmon. It designed the CSPDarknet53 by using the CSPNet to replace ResNet in YOLOv3 and used it as the backbone to extract features. It selected PANet to be used as a feature fusion network in lots of experiments, and SPP was also used to increase the receptive filed of network. It also used the Mish function as activate function of the backbone. Besides, it proposed to use the DIoU-NMS method to add the information of the center point distance to the Bounding Box screening process, and it also proposed new data augmentation method named Mosaic, etc. Compared with YOLOv3, the YOLOv4 has greatly improved. Based on YOLOv4, some improved methods are also proposed to improve accuracy or detection speed. Hu et al. proposed to add a spatial pyramid pooling network at the top and bottom of the backbone to increase the respective field to enhance the ability of feature extraction [18]. Besides, it also used the GIOU loss function instead of IOU loss function that was used in YOLOv4, to pay more attention to the non-overlapping area of the two kinds of boxes. It has higher accuracy than YOLOv4 in GWHD dataset. Cai et al. [19]proposed YOLOv4-5D that designed a new backbone CSPDarknet53_dcn for YOLOv4 by using deformable convolution at the last layer of CSPDarknet. Besides, to improve the detection accuracy of the smaller object, PAN++ is also designed, and two large-scale detection layers are also added. It has higher accuracy than YOLOv4 in KITTI dataset and BDD dataset.

The YOLO serial method is a typical object detection method based on deep learning and has also been widely used in insulator and defective insulator detection. Diana et al. [20] proposed real-time detection of insulators using UAV images based on YOLOv2 method. It also used some methods that are Gaussian noise, blurring, rotation, and scaling to argument dataset to avoid over-fitting. Han et al. [21] proposed a cascaded network model that contains two parts. The first part is used to locate the insulator string in an aerial image. The second part is used to detect the defective insulator from the located insulator string by the first part. Based on the cascaded network model, they also designed a three-structure spatial pyramid pooling model to improve the performance of locating insulator string in the first part, and used YOLOv3-tiny method to detect the insulator defective fault in the located insulator string region [22]. It has a faster detection speed than other methods and is more suitable for deploying on UAV. Liu et al. [23] proposed an improved YOLOv3 and applied it in defective insulator detection. It introduced the cross-stage partial dense network into the feature extract network to extract multi-scale features of insulator string. Besides, it also used complete intersection over union instead of the mean square error to improve loss function. It obtained better performance in

accuracy than YOLOv3 method on the modified Chinese Power Line Insulator Dataset. Some of the images captured by unmanned aerial vehicle are not clear. This will affect defective insulator detection. To solve the problem, it firstly designed a super-resolution convolution neural network to reconstruct the blurred insulator image, and secondly used the YOLOv3 method to detect the insulator strings and harmers [24]. A lightweight YOLO method named MTI-YOLO was proposed to realize the insulator detection under various background interferences [25], which was based on YOLOv3-tiny. It designed multi-scale feature detection headers and introduced the spatial pyramid pooling to the headers to improve the detection accuracy for different sizes of insulators. To extract more insulator semantic features, a multi-scale feature fusion method was proposed. Although it consumes longer detection times than YOLOv3-tiny, it has higher precision, recall and mean average precision than YOLOv3-tiny. Compared with YOLOv3, it has a faster detection speed and almost the same precision and mean average precision. Zhang et al. also designed a densely connected feature pyramid network to improve the YOLOv3 and used the improved method to detect normal insulator strings and defective insulators [26]. It has higher accuracy than the original YOLOv3. To further improve the insulator and insulator fault detection accuracy, some detection methods based on YOLO are also proposed [27, 28].

3. **Defective insulator based on improved YOLOv4.**

3.1. **Proposed dual-branch attention block.**

YOLOv4 uses a new backbone, CSP-Darknet53, the design of which is based on CSPNet. CSPDarknet53 increases the depth of the network to extract more features, but it also introduces more redundant semantic information. To suppress the redundant semantic information and reduce its influence on the accuracy of defective insulator detection, we introduce the channel attention mechanism into the depth network of the backbone. The channel attention mechanism can assign a larger weight to the channels that contain more important feature information relevant to the detection task. This makes the backbone extract more useful feature information from the vast pool of complex semantic information. Inspired by the efficient channel attention method [29], we have designed a dual-branch attention block that simulates the attention process of the human eye to extracting more relevant information from defective insulators. This is shown in Figure 1.

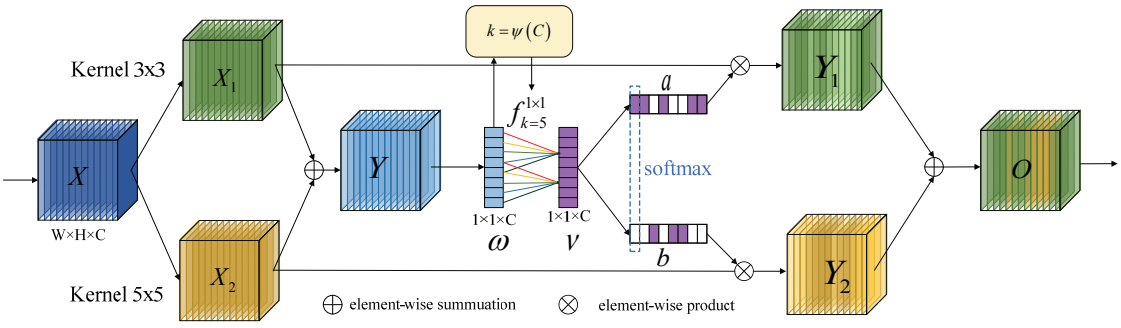


FIGURE 1. Proposed dual-branch attention block

The feature map firstly passes two branch and fuses in the end. In the proposed dual-branch attention block, expresses the input feature map, that is, the output of the previous feature extraction layer that is the CSP block in the YOLOv4 backbone. X_1 expresses the output feature map that X passes, a convolution layer with kernel size . X_2 expresses the

output feature map that X passes, a convolution layer with kernel size 5×5 and padding of 1. Therefore, we can obtain:

$$Y = X_1 + X_2 = f^{3\times3}(X) + f^{5\times5}(X) \tag{1}$$

where $f^{3\times3}()$ and $f^{5\times5}()$ are express convolution operations, with kernel size 3×3 padding of 1 and kernel size 5×5, and padding of 2, respectively. Next, we take channel-wise global average pooling operation on Y to obtain the channel level global statistics information. ω_c is the cth component of ω . It is expressed as follows:

$$\omega_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} Y_c(i,j) \tag{2}$$

where Y_c is feature map of cth channel, and H and W are the size of Y_c.

After obtaining the global information, we use a one-dimensional convolution operation to realize the cross-channel interaction that is useful for enhancing the information interaction between different channels. The range of channel interaction is determined by kernel size of one-dimensional convolution. We use the same method used in efficient channel attention to determine the kernel size of convolution. It can be expressed as follows [29]:

$$k = \psi(C) = \left| \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right|_{odd} \tag{3}$$

where C is the number of channels, and γ and b are pre-defined parameters (we set $\gamma = 2$ and $b = 1$). $|t|_{odd}$ expresses the nearest odd number of t. Based on (3), we can see that kernel size changes with a change in the number of channels. After determining the number of channel interactions, that is, the kernel size of a one-dimensional convolution, we can obtain the feature cross-channel interaction result. It is expressed as follows:

$$v = \sigma(\mathrm{C1D}_k(\omega)) \tag{4}$$

where $\mathrm{C1D}_k$ is the one-dimensional convolution with kernel size k, and $\sigma(\cdot)$is sigmoid function.

To obtain the weights of two branches, the feature map v s divided into two equal parts that are expressed as v_1 and v_2 , respectively. The weights of the two branches are expressed as $a = [a_1, a_2, \cdots, a_c]$ and $b = [b_1, b_2, \cdots, b_c]$, respectively. a_i and b_i are the ith channel weights of the two respective branches. They can be expressed as follows:

$$a_i = \frac{\exp(v_{1i})}{\exp(v_{1i}) + \exp(v_{2i})} \tag{5}$$

$$b_i = \frac{\exp(v_{2i})}{\exp(v_{1i}) + \exp(v_{2i})} \tag{6}$$

where v_{1i} and v_{2i} are the ith channel feature value of v_1 and v_2, respectively. The larger the channel feature value is, the larger the weight of channel is. Formula (5) and (6) are realized by using the softmax function. In the end, we get the output feature map O the ith channel feature map O_i is expressed as follows:

$$O_i = Y_{1i} + Y_{2i} = a_i \cdot X_{1i} + b_i \cdot X_{2i} \tag{7}$$

where X_{1i} and X_{2i} are the ith channel feature map of X_1 and X_2 in Figure 1, respectively. The larger the channel weight is, the larger ratio of the channel feature map is. This can make the output feature map contain more effective feature information of defective insulator. Using formula (7), we can fuse different feature maps originating from different respective fields so as to extract more effective feature information from redundant information.

3.2. **Proposed path aggregation network.** In YOLOv4 network, there are three different scale output feature maps in the backbone. PANet (path aggregation network) is used to fuse the features from the three different scales in YOLOv4 network. It firstly uses up-sampling to increase the scale of the shallow layer network to the scale of the deep layer network and uses concatenation and convolution operation to fuse the features of the shallow layer network and deep layer network. Secondly, it uses maxpooling to reduce the scale of the deep layer network to the scale of the shallow layer network and also uses concatenation and convolution operations to fuse the features of the deep layer network and shallow layer network. Based on the up-sampling and maxpooling, the output feature map of the shallow layer network obtains more fused features than the deep layer network. The scale of the output feature map for the shallow layer network is larger than the scale of the output feature map for the deep layer network. The larger-scale feature map is more suitable for detecting a smaller object, and the smaller feature map is more suitable for detecting a larger object. Therefore, the PANet used in YOLOv4 is better for detecting larger objects. In the defective insulator detection, the defective insulator is relatively small. Making the larger-scale feature map fuse more features can improve the defective insulator detection accuracy. Therefore, we propose an improved PANet to increase the fused features of larger scale feature map.

The improved PANet is shown in figure 2. Firstly, we use maxpooling to reduce the scale of the shallow layer network to the scale of the deep layer network and use concatenation and convolution operations to fuse the deep layer network and maxpooled feature features. Secondly, we also use the same method to fuse the deep layer network and deeper layer network. Thirdly, we use up-sampling to increase the scale of fused feature to the scale of deep layer network feature and use concatenation and convolution operations to fuse the features of shallow layer networks and deep layer network. Finally, we also use the same method to obtain the output feature of the shallow layer network. The structure of the improved PANet is the opposite of the original PANet. The improved PANet can obtain more features about the smaller object. For a more detailed description of the model, we suppose that the size of the input image is 416×416, so the sizes of feature maps for three branches of the backbone are 52×52, 26×26 and 13×13, respectively. We firstly use maxpooling to reduce the size of the feature map from 52×52 to 26×26 and concatenate the new 26×26 feature map with the original 26×26 feature map. The CBL block that contains a convolution layer, batch normalization, and Leaky ReLU activate function is used to fuse the concatenated 26×26 feature map. Secondly, we use maxpooling to reduce the size of the fused feature map from 26×26 to 13×13 and concatenate the new 13×13 feature map with the original 13×13 feature map. The CBL block also is used to fuse the concatenated 13×13 feature map, and the fused 13×13 feature map is used as the output feature map of the deep layer network. Thirdly, we use upsampling to increase the size of the fused feature map from 13×13 to 26×26, and concatenate the new 26×26 feature map with the fused 26×26 feature map that is obtained by upsampling, concatenating and fusing. The output of the fused 26×26 feature map is used as the output feature map of the intermediate layer network. In the end, we also use upsampling, concatenation,

and CBL modules to obtain the fused 52×52 feature map and use it as the output feature map of the deep layer network.

The output feature map of the deep layer network in the original PANet only contains the features that contains in the original 52×52 feature map, transmitted feature from 26×26 feature map by one time upsampling, and transmitted feature from 13×13 feature map by two times upsampling. In the improved PANet, the output feature map of deep layer network contains the original 52×52 feature map, transmitted feature from 26×26 feature map by one time upsampling, transmitted feature from 13×13 feature map by two times upsampling, transmitted feature from 26×26 feature map by one time maxpooling and transmitted feature from 52×52 feature map by two times upsampling and max-pooling. Therefore, the final 52×52 feature map contains more features in the improved PANet than the original PANet.

To further improve accuracy, we use the ASPP (atrous spatial pyramid pooling) structure instead of SPP (spatial pyramid pooling) structure in YOLOv4. The ASSP structure uses atrous convolution, instead of the max pooling operation used in SPP. It can increase the respective field and avoid losing information caused by the pooling operation. The relationship between the kernel size of normal convolution and atrous convolution is:

$$K = k + (k-1)(r-1) \tag{8}$$

where K is the kernel size of normal convolution, k is the kernel size of atrous convolution, and r is dilation rate. For the SPP network in the YOLOv4 method, the max pooling of kernel sizes are 3×3, $9 \times 9 and$ 13×13, respectively. Therefore, we respectively set the atrous convolution kernel sizes at 3×3, 5×5 and 7×7 with dilation rate 2, to make the ASPP network connect with the next layer. The ASPP network is shown in figure 2.

We introduce our proposed Dual-branch Attention Block (DBA) into the CSP block to construct the improved backbone. Besides, we use our proposed path aggregation network (PANet) and ASPP module to improve the neck part of YOLOv4. The improved YOLOv4 network for defective insulator detection is shown in Figure 2. Compared with YOLOv4, we introduce the DAB module, improved PANet module and ASPP module.

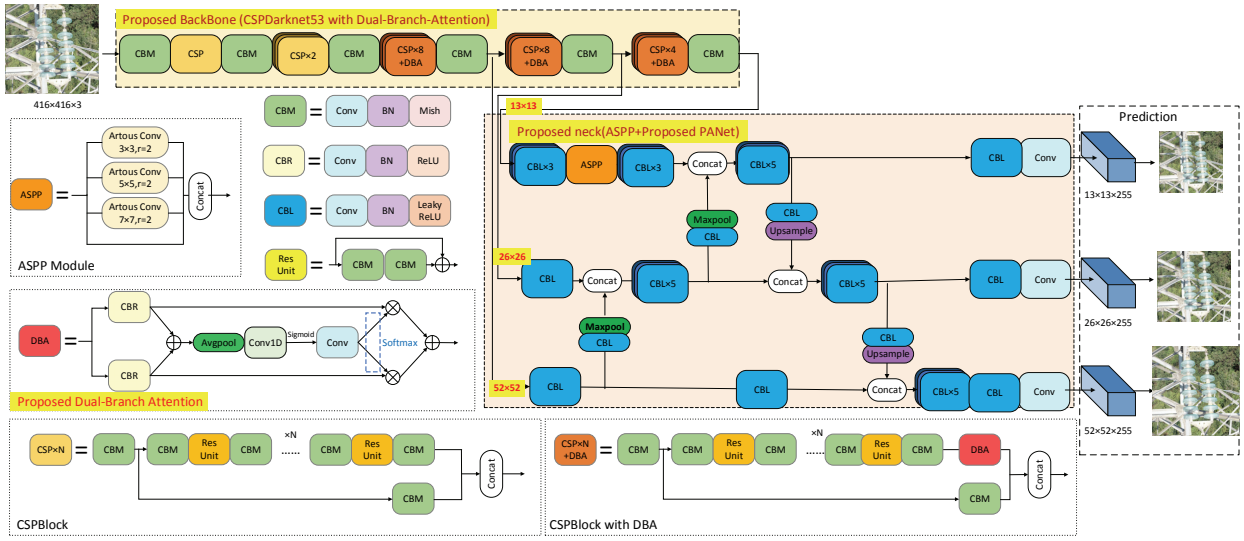


Figure 3

FIGURE 2. Proposed defective insulator detection network

3.3. **Proposed non-maximum suppression method.** The non-maximum suppression (NMS) method is used to select the optimal target box from the bounding boxes. In YOLOv4, it uses the DIoUNMS method, an improved NMS method for selecting the target box. DIoUNMS introduces the center distance between two bounding boxes into the threshold to avoid error suppression for different objects sharing a high degree of overlap. If the IOU (Intersection over Union) between the predicted boxes with low confidence and high confidence is larger than the threshold, the DIoUNMS method will directly delete the predicted box with low confidence. During real-world application, some predicted boxes with low confidence contain much useful information that could be used to improve detection accuracy. To make full use of a predicted box with low confidence and make the NMS dynamically adapt to high-density crowded scenes, a dynamic NMS method is proposed.

We propose an adaptive threshold control method for the high-density crowded scenes in NMS. To avoid error suppression for different objects with a high degree of overlap, the threshold should be larger when the number of objects and degree of overlap are larger. Nevertheless, the threshold should be smaller. Therefore, we introduce the number of objects and average IOU of the predicted box into the threshold and thus propose an adaptive threshold. The new threshold is expressed as follows:

$$T_{adap} = \frac{(D_{obj} + M_{iou} + T)}{3} \tag{9}$$

where T is the fixed threshold used in the original NMS, and $D_{obj} \in (0, 1)$ expresses same kind of object density. The larger the number of objects of the same kind is, the larger the number of predicted bounding boxes is, and the larger the value of D_{obj}. Therefore, we propose that D_{obj} can be expressed as follows:

$$D_{obj} = \frac{(D_{\max} - D_{\min})}{(N_{\max} - N_{\min})} \times N + D_{\min} \tag{10}$$

where $D_{\max}$ and $D_{\min}$ are the upper and lower limits of D_{obj}, respectively; and, $N_{\max}$ and$N_{\min}$ are the upper and lower limits of bounding boxes of defective insulators, respectively. In this paper, we set $D_{\max} = 0.04$, $D_{\min} = 0.01$, $N_{\max} = 20$ and $N_{\min} = 0$. N is the number of bounding boxs of defective insulators. $M_{iou} \in (0, 1)$ expresses the mean overlap level of the objects. The larger the overlap is, the larger the value of , which in turn is expressed as follows:

$$M_{iou} = \frac{1}{n} \sum_{i=1, i \neq best}^{n} IoU\left(b_{best}, b_i\right) \tag{11}$$

where $B = \{b_1, b_2, \cdots, b_n\}$ is the predicted box set of objects of the same kind, and b_{best} is the predicted box with the highest confidence score in B. In formula (9), we consider the density of object and overlap level in determining the NMS threshold in order to select a more optimal predicted box. In addition, in order to make full use of the predicted box with low confidence, we use the weighted operation to fuse these predicted boxes. The fused predicted box is the proposed optimal target box that is used to express the location of defective insulator. It is used in the Head network in Figure 3. It is expressed as follows:

$$B_{out} = \frac{\sum_{i=1}^{n} S_i \times IoU\left(b_i, b_{best}\right) \times b_i}{\sum_{i=1}^{n} S_i \times IoU\left(b_i, b_{best}\right)} \tag{12}$$

where S_i is confidence score of b_i box. b_i is expressed as follows:

$$b_i \in \{b_i | IoU\left(b_{best}, b_i\right) \geq T_{adap}\} \tag{13}$$

4. Simulation and discussion.

4.1. Datasets and simulation configuration.

We sourced 1545 insulator images that contain 45 defective insulator images for Power Supply and Internet. The number of defective insulator images is too small to train network. To increase the number of defective insulator images to avoid overfitting, we firstly use Photoshop to crop the insulator strings that contain defective insulators from the insulator images. Secondly, we use Photoshop to extract the cropped insulator strings from background by background color eraser tools. In the end, we use Photoshop to fuse the extracted insulator with different background to generate new defective insulator images.

The generation of a new defective insulator image is shown in Figure 3. The Figure 3(a) image is the original image that contains a defective insulator. The Figure 3(b) image is the image cropped from Figure 3(a). The Figure 3(c) image is the extracted image with transparent background from Figure 3(b). The Figure 3(d) image is the insulator image that is used to be fused with Figure 3(c) image. The Figure 3(e) is the fused image. Compared with the original image Figure 3(a), the background of the fused image is completely different. To further compare the difference between the fused image and the original image, we also crop the defective insulator string from the fused image (e) to obtain the enlarged insulator string image Figure 3(f). In Figure 3(b), the nearby background of defective insulators is a leaf. In Figure 3(f), the nearby background of defective insulators is land. This shows that the nearby backgrounds of defective insulators are also different between Figure 3(b) and Figure 3(f). The defective insulators are more difficult to recognize by our eyes in Figure 3(f) than Figure 3(b). The background has a large impact on defective insulator detection. Based on this method, we select 30 defective insulator images as original training images and 15 defective insulator images as original test images. We use each original defective insulator image to generate 20 images with different backgrounds. The training dataset contains 1000 normal images and 600 defective insulator images. The test dataset contains 500 normal images and 300 defective insulator images.

4.2. Performance of proposed model.

In this paper, we propose a modification to the original YOLOv4 by introducing our proposed Dual-Branch Attention into the backbone of YOLOv4 to suppress redundant semantic information and reduce the influence of redundant information on detection accuracy, designing a path aggregation network to obtain more location information of defective insulator and proposing synthetic Non-Maximum suppression method to select a more optimal predicted box. For the sake of simplicity, we name the Dual-Branch Attention as DBA, path aggregation network as PAN, and synthetic Non-Maximum suppression as Syn. The defective insulator detection average precisions for YOLOv4 based on different proposed methods are shown in table 1. The YOLOv4 based on DBA has the highest mean average precision, followed by YOLOv4 based on PAN and YOLOv4 based on Syn. Compared with YOLOv4, the mean

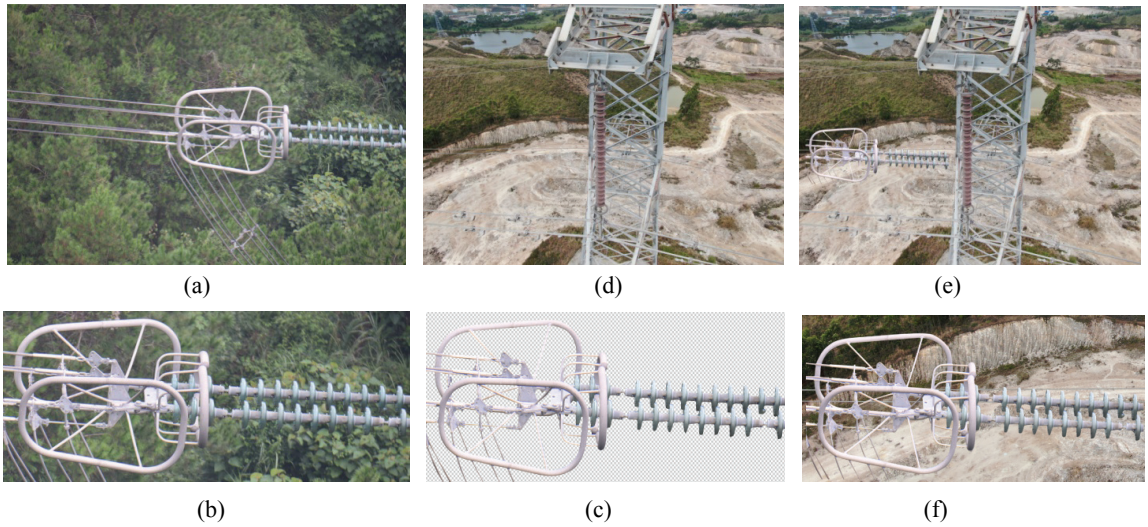


FIGURE 3. Generation of new defective insulator image. (a) original image with defective insulators;(b) cropped image from (a); (c) extracted image with transparent background from (b); (d) image with different background; (e) image obtained by fusing (c) and (d); (f) enlarged defective insulator string from (e)

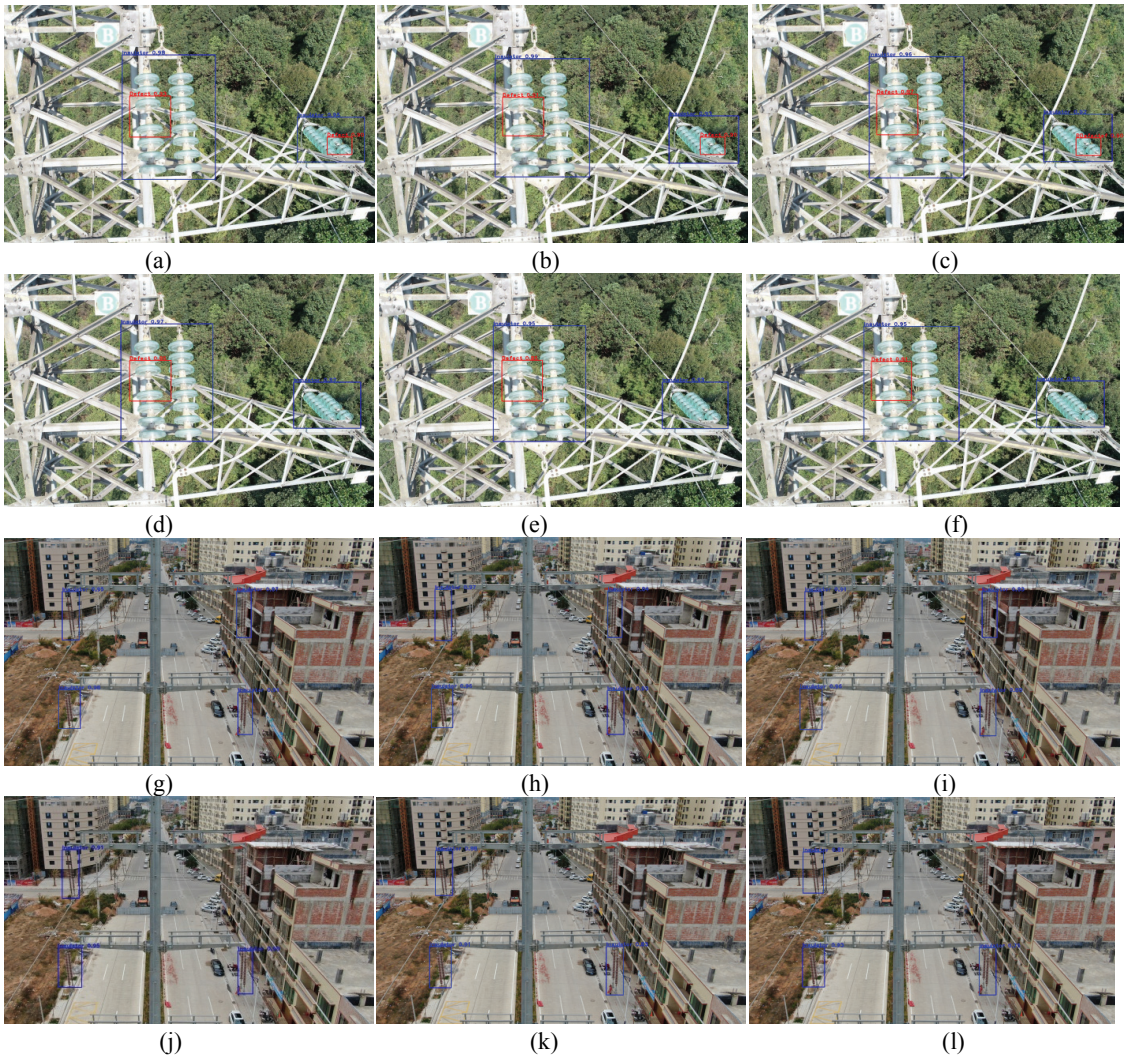
average precision of YOLOv4 based on PAN, YOLOv4 based on DBA, and YOLOv4 based on Syn are 2.9%, 3.3% and 1.2% higher, respectively. These results indicate that our proposed PAN, DBA, and Syn are effective.

TABLE 1. Accuracies of different methods on office-31 dataset.

	PAN	DBA	Syn	mAP(%)
YOLOV4				79.2
YOLOV4	√			82.1
YOLOV4		√		82.5
YOLOV4			√	80.4

We show some visual detection results with different models in Figure 4. Figure 4(a)-(h) show the detection results of two insulator strings and two defective insulators using model [30], model [12], our model, YOLOv4, InsuDet [31] and YOLOv3, respectively. Figure 4(g)-(l) show the detection results of four insulator strings without defective insulators using model [30], model [12], our model, YOLOv4, InsuDet, and YOLOv3, respectively. In figure 4(a)-(h), the model [30], model [12], and our model successfully detect all insulator strings and defective insulators, and the YOLOv4, InsuDet, and YOLOv3 detect two insulator strings and one defective insulator. The YOLOv4, InsuDet, and YOLOv3 do not detect the defective insulator that locates at the right of the image. The feature of the right defective insulator is not obvious, and its size is relatively smaller, so it is difficult to detect. The larger scale detection branch of our model based on improved PANet and ASPP nework obtains more fused information, so it has better accuracy in the detect smaller defective insulator. In Figure 4(g)-(l), the model [30], model [12] and our model also successfully detect all four insulator strings, and the YOLOv4, InsuDet, and YOLOv3 detect three insulator strings. The YOLOv4, InsuDet, and YOLOv3 do not detect the insulator string that locates at the top right of the image. The color of the undetected insulator string is very close to the color of the surrounding background

(building), so it is difficult to detect. We propose a new dual-branch attention block and introduce it into the backbone to make the model pay more attention to the effective feature. This reduces the background interference. Although the detection results are influenced by background and the size of the object, the model [30], model [12], and our model successfully detects all objects. This shows that model [30], model [12], and our model has better performance in accuracy than others.



1

FIGURE 4. Detection results with different models. (a)-(f) detection results using model[30], model[12], our model, YOLOv4, InsuDet and YOLOv3, respectively. (g)-(l) detection results using model [30], model [12], our model, YOLOv4, InsuDet and YOLOv3, respectively. Blue box: Bounding box of detected insulator string. Red box: bounding box of detected defective insulator.

To quantitatively analyze the performance of our model, we use all test images to test the performance of different models. The results of our proposed model and other model are shown in Table 2. The model [30] has the highest mAP and Recall, followed by model [12] and our model. The model [12], and model [30] are based on R-CNN and belong to two-stage detection method. The YOLOv3, InsuDet, YOLOv4, and our model are based on YOLO and belong to one-stage detection method. The two-stage model focus on accuracy, and the one-stage method focuses on detection speed. Therefore, model [12],

and model [30] have better performance in mAP and Recall for detecting insulator strings and defective insulators than on-stage models that are YOLOv3, InsuDet, YOLOv4, and our model. Compared with YOLOv3, InsuDet, and YOLOv4 methods, the mean average precisions of our proposed method are 14.4%, 7.1%, and 6.8% higher, and the recalls of our proposed model are 12.7%, 6% and 4.2% higher, respectively. Compared with model [30] and model [12], the mean average precisions of our proposed model are 6.5% and 5.4% lower, and the recalls of our proposed model are 6.2% and 4.4% lower, respectively. Therefore, based on the analysis of Table2, although our proposed model has lower performance in mAP and Recal than model [12] and model [30], it has higher performance in mAP and Recall than YOLOv3, InsuDet, and YOLOv4.

Table 2. Different models on mean average precision, Recall.

Model	mAP(%)	Recall(%)
Model[30]	92.5	96.7
Model [12]	91.4	94.9
YOLOv3	72.4	77.8
InsuDet	78.9	84.5
YOLOv4	79.2	86.3
Our model	86.0	91.5

We also compare the detection speed of different models. The average consuming times of detecting one image for different models are shown in Table 3. The model [30] and model [12] belong to two-stage detection methods, so they have the longer detection time. On the contrary, YOLOv3, InsuDet, YOLOv4 and our model belong to one-stage detection methods, so they have shorter detection time. Compared with model [30] and model [12], detection speeds of our proposed model are about 4.3 times and 9.7 times, respectively. Compare with YOLOv3, InsuDet and YOLOv4, the detection times of our proposed model are about 4.3ms, 5.5ms and 2.2ms longer for detecting one image, respectively. Our model introduce dual-branch attention model into backbone, it increases complexity of network and consumes more detection time.

Based on Table 2 and Table 3, compared with two-stage models (model [30] and model [12]), our model has a faster detection speed and lower mean average precision and Recall. The mean average precision and Recall of our model are about 6% lower than model [30] and model [12], but the detection speeds of our model are about nine times and four times than model [12] and model [30], respectively. Compared with one-stage models (YOLOv3, InsuDet, and YOLOv4), our model has a slower detection speed, and higher detection mean average precision and Recall. Although our model has a slower detection speed than YOLOv3, InsuDet, and YOLOv4 models, the difference in detection speed has little influence on real-time detection. Our method has higher detection accuracy with acceptable detection speed, indicating that our proposed method is more suitable for the real-time detection of defective insulators.

TABLE 3. Different models on detection speed.

Model	Average detection time(ms/image)
Model[30]	138.7
Model [12]	314.6
YOLOv3	28.2
InsuDet	27.0
YOLOv4	30.3
Our model	32.5

5. **Conclusions.** This paper proposes a new defective insulator detection method that offers higher accuracy and a more acceptable detection speed. In addition, it also details the construction of a new defective insulator dataset, whereby Photoshop is used to extract defective insulators and fuse them into different backgrounds. In the backbone, it introduces a new dual-branch attention block in order to extract more effective defective insulator feature information from the redundant semantic information. The novel path aggregation network is designed to fuse more effective information relating to defective insulators. In addition, a dynamic non-maximum suppression method for improving defective insulator detection accuracy has been outlined. Compared with the defective insulator detection method that has the highest mean average precision in our simulations, the mean average precision of our method is 6.5% lower, but the detection speed of our method is 4.3 times faster. The increased detection speed is larger than the reduced mean average precision. Compared with the defective insulator detection method that has the fastest speed in our simulations, the detection speed of our method is 5.5ms lower, but mean average precision of our method is 7.1% higher. Our proposed method strikes a careful trade-off between detection speed and precision.

Although the improved YOLOv4 has higher precision with acceptable detection speed, it cannot be employed on lightweight equipment that unmanned aerial vehicles can carry. In our feature work, we will consider how to reduce the complexity of improved YOLOv4 to operate it on lightweight equipment. Therefore, the defective insulator detection can be realized live without transmitting the image to the remote server.

REFERENCES

[1] J.-Y. Liang, Y.-L. Wang, J.-B. Li, and X.-N. Feng, "Research on Insulator Flashover Criterion Based on Leading Progression Model," *Journal of Northeast Electric Power University*, vol. 41, no. 5, pp. 74-79, 2021.

[2] Y.-J. Liu, X.-Y. Zhang, J.-S. Guo, and Z. Meng, "Study on the Reliability of Electric Transmission Line under Ice Load and Wind Load," *Journal of Northeast Electric Power University*, vol. 40, no. 5, pp. 63-68, 2020.

[3] M.-E. Wu, J.-H. Syu, and C.-M. Chen, "Kelly-Based Options Trading Strategies on Settlement Date via Supervised Learning Algorithms," *Computational Economics*, vol. 59, no. 4, pp. 1627-1644, 2022.

[4] S. Kumar, A. Damaraju, A. Kumar, S. Kumari, and C.-M. Chen, "LSTM Network for Transportation Mode Detection," *Journal of Internet Technology*, vol. 22, no. 4, pp. 891-902, 2021.

[5] E.-K. Wang, X. Zhang, F. Wang, T.-Y. Wu, and C.-M. Chen, "Multilayer dense attention model for image caption," *IEEE Access*, vol. 7, pp. 66358-66368, 2019.

[6] L.-Z. Yu, B. Xiao, and L.-G. Sun, "Conditional Depth Convolution Generation of Confrontation Network Method for Scenery Output Scenario Generation," *Journal of Northeast Electric Power University*, vol. 41, no. 6, pp. 90-99, 2021.

[7] B. Alexey, W. Chienyao, and L.-H. Mark, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint*, 2022. [Online]. https://arxiv.org/abs/2004.10934

[8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2016, pp. 779-788.

[9] X.Lei and Z. Sui, "Intelligent fault detection of high voltage line based on the Faster R-CNN," *Measurement*, vol. 138, pp. 379-385, 2019.

[10] Z.-B. Zhao, Z. Zhen, L. Zhang, Y.-C. Qi, Y.-H. Kong, and K. Zhang, "Insulator Detection Method in Inspection Image Based on Improved Faster R-CNN," *Energies*, vol. 12, no. 7, pp. 1-15, 2019.

[11] Z.-N. Ling, D.-Z. Zhang, R.-C. Qiu, Z.-J. Jin, Y. Zhang, X. He and H. Liu, "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.

[12] X. Tao, D. Zhang, Z. Wang, X. Liu, H. Zhang and D. Xu, "Detection of Power Line Insulator Defects Using Aerial Images Analyzed With Convolutional Neural Networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 4, pp. 1486-1498, 2020.

[13] X. Liu, X. Miao, H. Jiang, and J. Chen, "Box-Point Detector: A Diagnosis Method for Insulator Faults in Power Lines Using Aerial Images and Convolutional Neural Networks," *IEEE Transactions on Power Delivery*, vol. 36, no. 6, 3768-3773, 2021

[14] W. Zhao, M. Xu, X. Cheng, and 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, pp. 1-8, 2021

[15] H. Ma, Y. Liu, Y. Ren, and J. Yu, "Detection of Collapsed Buildings in Post-Earthquake Remote Sensing Images Based on the Improved YOLOv3," *Remote Sensing*, vol. 12, no. 1, 44, 2020.

[16] Y. Yang and D. Hongmin, "GC-YOLOv3: you only look once with global context block," *Electronics*, vol. 9, no. 8, 1235, 2020.

[17] C. Liu, Y. Wu, J. Liu, and Z. Sun, "Improved YOLOv3 Network for Insulator Detection in Aerial Images with Diverse Background Interference," *Electronics*, Vol. 10, no. 7, 771, 2021.

[18] X. Hu, Y. Liu, Z. Zhao, J. Liu, X. Yang, C. Sun, S. Chen, and B. Lin, "Real-time detection of uneaten feed pellets in underwater images for aquaculture using an improved YOLO-V4 network," *Computers and Electronics in Agriculture*, vol. 185, 106135, 2021.

[19] Y. Cai, T. Luan, H. Gao, H. Wang, and Z. Li, "YOLOv4-5D: An Effective and Efficient Object Detector for Autonomous Driving," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-13, 2021.

[20] S. Diana, P. Damira, B. Mehdi, and J. Alex, "IN-YOLO: Real-Time Detection of Outdoor High Voltage Insulators Using UAV Imaging," *IEEE Transactions on Power Delivery*, vol. 35, no. 3, pp. 1599-1601, 2020.

[21] J. Han, Z. Yang, Q. Zhang, C. Chen, and R. Chen, "A Method of Insulator Faults Detection in Aerial Images for High-Voltage Transmission Lines Inspection," *Applied Sciences*, vol. 9, no. 10, 2009, 2019.

[22] J. Han, Z. Yang, H. Xu, G. Hu, C. Zhang, H.-Li, S. Lai, and H. Zeng, "Search Like an Eagle: A Cascaded Model for Insulator Missing Faults Detection in Aerial Images," *Energies*, vol. 13, no. 3, 713, 2020.

[23] C. Liu, Y. Wu, J. Liu, Z. Sun, and H. Xu, "Insulator Faults Detection in Aerial Images from High-Voltage Transmission Lines Based on Deep Learning Model," *Applied Sciences*, vol. 11, no. 10, 4647, 2021.

[24] H. Chen, Z. He, B. Shi, and T. Zhong, "Research on Recognition Method of Electrical Components Based on YOLOv3," *IEEE Access*, vol. 7, pp. 157818-157829, 2019.

[25] C.-Y,Liu, W. Yiquan, L. Jingying, and H. Jiaming, "MTI-YOLO: A Light-Weight and Real-Time Deep Neural Network for Insulator Detection in Complex Aerial Images," *Energies*, vol. 14, no. 5, 1246, 2021.

[26] X. Zhang, Y. Zhang, J. Liu, and C. Zhang, "InsuDet: A Fault Detection Method for Insulators of Overhead Transmission Lines Using Convolutional Neural Networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-12, 2021.

[27] Y.-J. Wang, P.-P. Cao, X. S. Wang, and X.-Y. Yan, "Research on Insulator Self Explosion Detection Method Based on Deep Learning," *Journal of Northeast Electric Power University*, vol. 40, no. 3, pp. 33-40, 2020.

[28] C. Liu, Y. Wu, J. Liu, and Z. Sun, "Improved YOLOv3 Network for Insulator Detection in Aerial Images with Diverse Background Interference," *Electronics*, vol. 10, no. 7, 771, 2021.

[29] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, 2020, pp. 11531-11539.

[30] W. Zhao, M. Xu, X. Cheng, and 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, pp. 1-8, 2021

[31] X. Zhang, Y. Zhang, J. Liu, and C. Zhang, "InsuDet: A Fault Detection Method for Insulators of Overhead Transmission Lines Using Convolutional Neural Networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-12, 2021.