*Journal of Network Intelligence*
**Taiwan Ubiquitous Information**
©2021 ISSN 2414-8105 (Online)
Volume **6**, Number **4**, November **2021**

# Detection of Insulator Defects Based on Improved YOLOv3

Ren-Jie Song

Department of Computer Science
Northeast Electric Power University
No. 169 Changchun Road, Jilin, Jilin, China
1939811347@qq.com

Dong-He Jin*

Department of Computer Science
Northeast Electric Power University
No. 169 Changchun Road, Jilin, Jilin, China
Corresponding Author: 506433601@qq.com

Khishiguren Dovagdorj

Department of Electrical and Computer Engineering
Chungbuk National University
Cheongju 28644, Republic of Korea
khishigsurend@chungbuk.ac.kr

Received May 2021; revised August 2021

Abstract. *At present, UAV is widely used in insulator defects detection of transmission line. In order to effectively detect insulator defects in aerial images in real time, an insulator defects detection method based on improved yolov3 is proposed. According to the definition of receptive field in convolutional neural network, a new yolov3 targets detection network is constructed. Then the depthwise separable convolution and anti residual block are used to reduce the size of the model and improve the detection speed. Finally through the cascaded structure of two object detection models, the insulator defects detection problem is transformed into a two-stage object detection problem. The experimental analysis shows that this method can effectively locate the defects in insulators in real time, and a good compromise is achieved between the speed and accuracy of insulator defects detection. It has good practicability in insulator defects detection based on UAV.*
**Keywords:** insulator defects, object detection; YOLOv3, depthwise separable convolution, anti residual block

1. **Introduction.** Due to their close relation to daily life and the national economy, power grids have become particularly important [1-2]. Insulators are essential equipment in high-voltage power transmission system [3]. They are made of high-strength insulating materials such as ceramics, tempered glass or synthetic rubber, which play a role of support and insulation in the operation of power grids [4]. With the rapid construction of China's power grids and the growth of the number of insulators, power accidents caused by insulator defects are also increasing. Defects of the insulator mainly come from its poor working environment [5]. Snow, haze and bird droppings in the field will result in flashover of the external insulation of the insulators, lightning may cause electrical breakdown, and storm and external force damage will cause damage to the mechanical structure of the

insulators [6]. Once the insulator is defective, it will greatly improve the risk probability of failure of the entire high voltage line and pose a great threat to the safe operation of the power grids. Therefore, ensuring the effective working state of insulator is particularly important to maintain the safe and stable operation of power grids [7]. Transmission line inspection is an essential means to ensure the normal operation of electrical equipment and maintain the safe and stable operation of power grids. The traditional transmission line inspection adopts manual patrols or manned helicopter inspection to inspect the relevant equipment of the transmission line, which is inefficient and costly. The review of staff based on subjective experience also has the problem of insufficient accuracy. In recent years, with the development and progress of science and technology, more and more enterprises and institutions began to implement the "Intelligent patrol inspection" of the combination of UAV and artificial intelligence technology [8]. Intelligent patrol inspection refers to the patrol inspection technology that uses UAV to collect data and images of power lines, and then uses computer vision technology to analyze the collected image information instead of human beings, so as to finally realize digital and unmanned operation process. Intelligent patrol inspection has the characteristics of accuracy, safety, efficiency and economy. It is the advance direction of Power Patrol Inspection Technology in the future. In practical applications, the images captured by UAV usually contain complex geological environment backgrounds such as towers, mountains, rivers, grasslands and farmland. These factors are easy to lead to sinful defect detection.

Wang et al. [9] used YOLO to find the insulators in the images, and then the insulator masks are extracted based on the salience detection. Finally, the horizontal projection method is used to locate the insulator defects' location. However, this method of locating the defect position in the horizontal projection curve is sensitive to noise, which will affect the detection effect. Guo et al. [10] used the Faster-RCNN object detection framework to detect the insulator strings on the images taken by the drone, and then converted the RGB images of the detected insulator strings area to the HIS space, and performed threshold segmentation on the H channel and the S channel. This method extracted the insulator strings from the background area, and finally determined whether the insulator string had defects according to the distance between the insulator sheets. Although this method can detect insulator strings of high accuracy, the two-stage objects detection framework Faster-RCNN requires higher hardware resources, and compared with one-stage object detection algorithms such as YOLO and SSD, it is slower in detection and cannot be embedded into UAV. Zheng et al. [11] used the improved YOLOv3 method to detect insulators and other power equipment in infrared images, but could not detect defects in the equipment. Zhao et al. [12] used OAD-BSPK to locate multiple insulators with different angles in complex aerial images. The above methods usually locate the insulators first, and then extract the characteristics of the insulators and determine their state by using the traditional image processing method or shallow learning method. These methods usually have certain requirements for the size, shape, angle and background of insulators, so they are not appropriate for practical application. To solve the above problems, a complete solution of insulator defects detection based on improved YOLOv3 algorithm is proposed. In order to better to detect insulator defects, cascade architecture is adopted. Firstly, the insulators are detected, and the detected insulators are used as input to detect the defects in the insulators. In order to compensate for the increased amount of computation caused by the cascade architecture, the detection speed is improved by optimizing the network structure. This method is more universal and can locate insulator defects quickly and accurately.

## 2. **Insulator defects detection.**

2.1. **YOLO object detection algorithm.** The two-stage object detection algorithm does not share the convolution parameters between the candidate region recommendation network and the detection network. For example, the Faster R-CNN [13] algorithm first uses RPN network to generate candidate regions, and then the feature extraction net is used to classify the objects in the candidate region and return to the bounding box. This algorithm has the disadvantage of slow detection speed.

Two-stage object detection frameworks are generally slow in detection. In order to improve the speed of object detection, Redmon et al. [14] proposed an algorithm called YOLOv1, which uses features to extract the last layer of the network's feature map to classify the target category and regression the bounding box. YOLOv1 is not good at detecting objects and small objects that are close to each other. This is because YOLOv1 only predicts in a feature map of one scale and only predicts two boundary boxes in one grid. Later, a series of improved YOLO versions used a deeper and more powerful feature extraction network, and used the strategy of setting Anchor Box (a priori box) in Faster R-CNN, using a full convolutional network, and passing the prediction of the prior box Offset to reduce the difficulty of network training. YOLOv2 [15] overcomes the shortcomings of YOLOV1's low recall rate and high positioning error rate, while improving the accuracy of object classification. YOLOv3 [16] further improves the accuracy rate than YOLOv2. In terms of network structure, YOLOv3 combines the multi-branch convolution of GoogLeNet [17] and the direct connection method of ResNet [18]. Although YOLOv3 has high detection accuracy, it is also much faster than the two-stage object detection framework such as Faster-RCNN. However, due to the deep network structure of darknet-53, its real-time performance on low-performance embedded devices is not ideal. Therefore, this paper proposes a lightweight YOLOv3 model, and refers to the RFB Net that simulates the receptive field of human vision to enhance the feature extraction ability of the network, which makes up for the lack of feature extraction ability after the network depth is reduced. In order to improve the real-time performance when the accuracy of model detection is slightly reduced but still meets the requirements. Figure 1 shows the network structure of YOLOv3.

2.2. **Improving the network structure of YOLOv3.** Although YOLOv3 has high detection accuracy, it is also much faster than the two-stage object detection frameworks such as Faster R-CNN. However, since the darknet-53 it uses has a relatively deep network structure, the real-time performance on low-performance devices or PCs is not ideal.

In this paper, a new feature extraction network is used to replace the feature extraction network darknet-53 of YOLOv3. The improved algorithm reduces the regression scale of the detection algorithm and the number of layers of the backbone network. The deep separable convolution is used to reduce the amount of network calculation. The feature maps of different scales have different receptive fields. Therefore, the receptive fields are improved by adding multi-scale branches to the three scale feature extraction layers of the backbone network. Firstly, each branch is down sampled, and the feature map with a smaller scale has a larger receptive field. Then, the feature extraction is carried out on the feature map of this scale, and the feature fusion is carried out with the feature map of the main branch after up sampling by the nearest neighbor interpolation method. This structure can better extract and integrate local and global information of different scales, improve the ability of shallow network to extract semantic information and improve the ability of target feature detection. The improved network structure of YOLOv3 is shown in Figure 2.

In the figure2, Conv3 is a convolution kernel with a size of $3\times3$, Conv1 is a convolution kernel with a size of $1\times1$, RFB is an RFB layer, and ARB is an anti residual block that

| | Type | Filter | | Feature map | |
|---|---|---|---|---|---|
| | Convolutional | 32 | 3×3 | 416×416 | |
| | Convolutional | 32 | 3×3/2 | 208×208 | |
| 1× | Convolutional | 32 | 1×1 | | |
| | Convolutional | 64 | 3×3 | | |
| | Residual | | | 208×208 | |
| | Convolutional | 128 | 3×3/2 | 104×104 | |
| 2× | Convolutional | 64 | 1×1 | | |
| | Convolutional | 128 | 3×3 | | |
| | Residual | | | 104×104 | |
| | Convolutional | 256 | 3×3/2 | 52×52 | |
| 8× | Convolutional | 128 | 1×1 | | |
| | Convolutional | 256 | 3×3 | | |
| | Residual | | | 52×52 | predict 3 |
| | Convolutional | 512 | 3×3/2 | 26×26 | |
| 8× | Convolutional | 256 | 1×1 | | |
| | Convolutional | 512 | 3×3 | | |
| | Residual | | | 26×26 | Predict 2 |
| | Convolutional | 1024 | 3×3/2 | 13×13 | |
| 4× | Convolutional | 512 | 1×1 | | |
| | Convolutional | 1024 | 3×3 | | |
| | Residual | | | 13×13 | Predict 1 |
| | Avgpool | | Global | | |
| | Connected | | 1000 | | |
| | Softmax | | | | |

FIGURE 1. Structure of YOLOv3

uses Depthwise separable convolution, the ARB with step size of 2 is used to replace the max pooling for feature down sampling. And perform up-sampling in the 13×13 feature layer, enlarge the feature map to 26×26, and superimpose it with the feature map of the previous layer with a size of 26×26, finally the prediction is carried out on two scales of 13×13 and 26×26.

2.3. **Improved Convolutional layer.** The improved detection algorithm uses a Depthwise separable convolution composed of Depthwise (DW) convolution and Pointwise (PW) convolution [19] instead of conventional convolution operations. The structure is shown in Figure 3. Different from the conventional convolution operation, one convolution kernel is only responsible for one channel in the DW convolution machine, and one channel is convolved by only one convolution kernel. The number of convolution kernels is equal to the output dimension of the previous layer feature map, that is, the input feature map of this layer corresponds to the convolution kernel one by one. Therefore, the dimension of the feature maps cannot be expanded, and the feature information of different channels at the same spatial position is not effectively used. Therefore, we add PW convolution to merge these feature maps to generate new maps. PW convolution uses $1 \times 1$ convolution kernel to weight sum the feature maps of the previous layer in the depth direction and generate a new feature map.

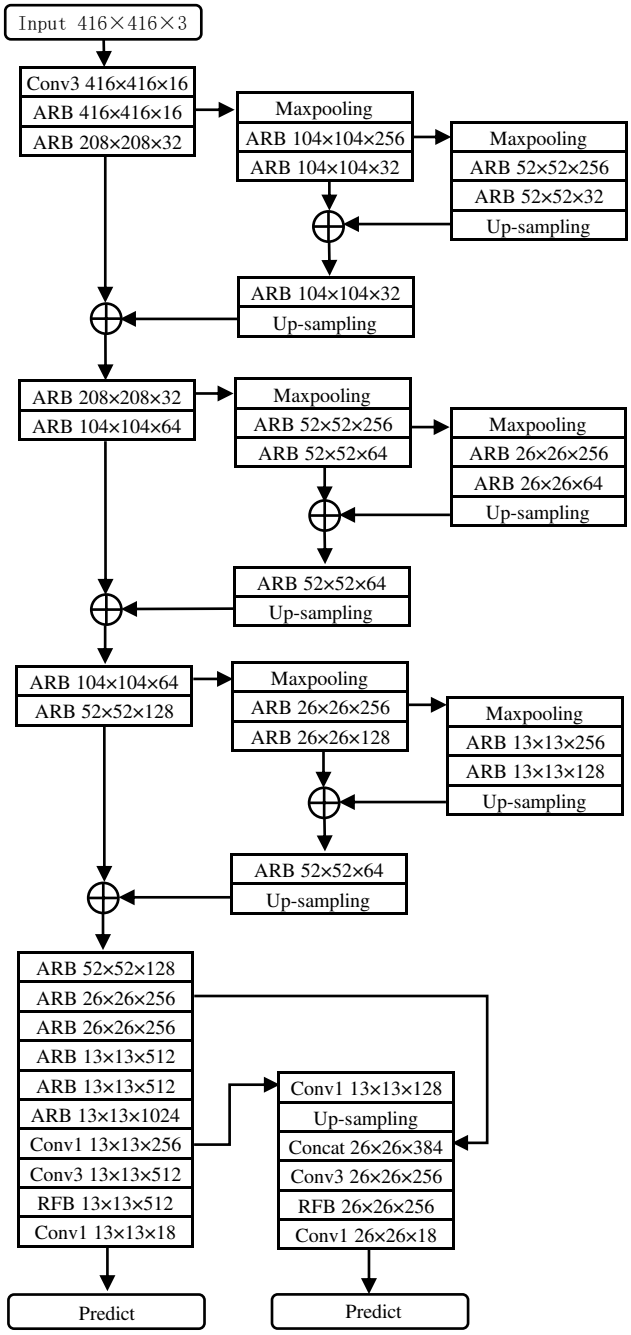


FIGURE 2. Structure of improved YOLOv3

Suppose the size of the original images is $D_F \times D_F$, the dimension is $M$, the size of the Convolution kernel is $D_K \times D_K$, and the number of channels is $N$. The calculation amount of ordinary convolution operation is:

$$C_n = D_F \times D_F \times M \times N \times D_K \times D_K \tag{1}$$

The calculation amount of the Depthwise separable convolution is:

$$C_d = D_F \times D_F \times M \times D_K \times D_K + D_F \times D_F \times M \times N \tag{2}$$

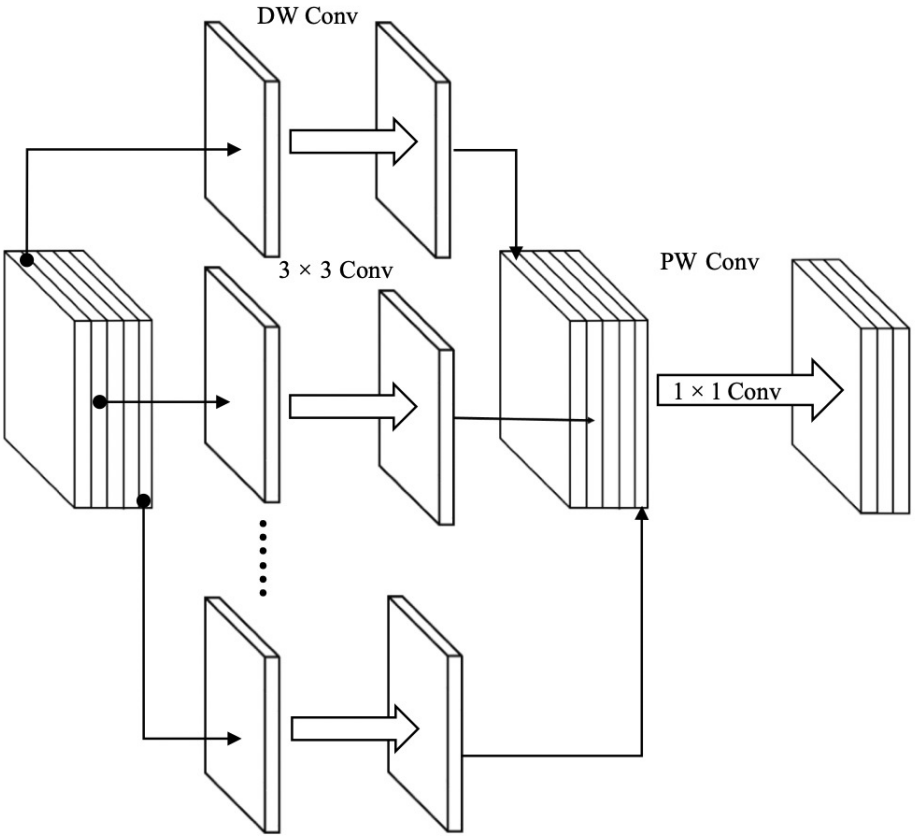


FIGURE 3. Depthwise separable convolution

The ratio of the calculation amount of the Depthwise separable convolution to the calculation amount of the ordinary convolution is:

$$\frac{C_n}{C_d} = \frac{1}{N} + \frac{1}{D_K^2} \tag{3}$$

The DW convolution can effectively reduce the computational complexity and the size of the model, but it can not solve the problem that the gradient disappears with the increase of convolution layers. Ordinary residual structure uses 1×1 convolution operation to map high dimensional space to low-dimensional space, which will compress feature map and damage feature expression. Therefore, we used the anti-residual block. First, the number of channels is expanded to 4 times through 1×1 convolution, and then 3×3 DW convolution is used to extract high-dimensional spatial features, and finally through 1×1 The PW convolutional layer maps the result to the new feature space. Figure 4 is the internal network structure diagram of the anti-residual structure used in this paper. In order to prevent gradient explosion, gradient disappearance and speed up the network convergence, the output value after each reel operation is normalized, which makes it easier and more stable to train the neural network model. According to the definition of the receptive field in the convolutional neural network, the receptive field module is added to the feature extraction network, which is called the RFB layer in this paper. The RFB layer combines the ideas of multi-branch convolution (Inception) and hollow convolution to simulate human visual perception as much as possible [20], strengthen the feature extraction ability when using a lightweight backbone network. The structure is shown in Figure 5.

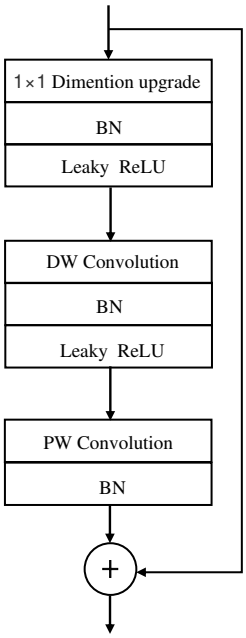


FIGURE 4. Internal structure of Anti Residual Block

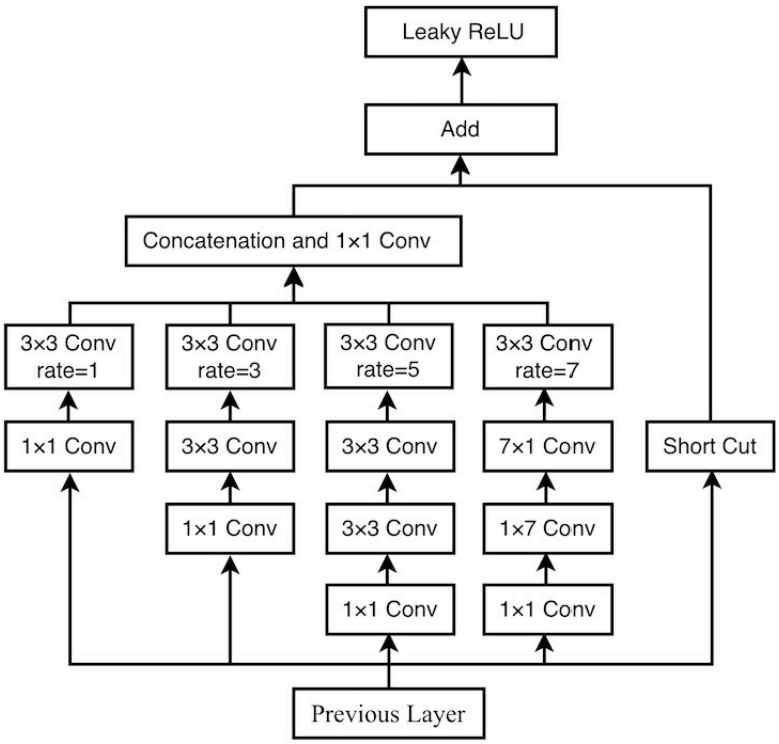


FIGURE 5. Structure of RFB layer

RFB layer adopts multi branch structure, and the convolution kernel size of each branch is different. In order to decrease the training parameters and network training complexity, each branch adopts the bottleneck structure.The shortcut structure integrating the features of the previous layer and the features of the current layer is used to solve the problem of gradient divergence when increasing the network depth, To reducing the amount of network parameters and calculations, replaced the $5 \times 5$ convolution with two $3 \times 3$

convolutions, and used $1 \times 7$ and $7 \times 1$ asymmetric convolution replaced $7 \times 7$ convolution. The dilated convolution adds a parameter hole rate on the basis of the original convolution, which is recorded as $R$. This parameter controls the degree of expansion of the convolution kernel, and fills the unoccupied area in the original convolution kernel with 0. The effective height and width of the dilated convolution kernel are as following:

$$H = F_H + (F_H - 1)(R - 1) \tag{4}$$

$$W = F_W + (F_W - 1)(R - 1) \tag{5}$$

The height and width of the original convolution kernel are represented by $F_H$ and $F_W$. Specifically, the ordinary convolution operation is to multiply the adjacent pixels of the feature map, while the dilated convolution operation is to multiply the pixels with a fixed interval $R$. Without increasing the number of extra calculations, the receptive field is improved.

2.4. **Selection of Anchor-box.** The original YOLO algorithm needs to manually select the size of a priori box, but the size of the manually selected a priori box does not match the size of the target object well. In order to optimize the regression speed of the training object's bounding box, the size of the prior box should be as close as possible to the size of the detected object. Therefore, we need to get the size of the prior box by K-means clustering algorithm on the boundary box of the training set. In the calculation process of K-means algorithm, the distance between objects is usually the Euclidean distance. However, it only cares about the distance between the centers of objects, and cannot describe the overlapping relationship between objects of different sizes. Therefore, the IOU is used in this paper. IOU is the ratio of the intersection and union of two object frames. The formula for calculating the distance of IOU is as follows following:

$$IoU = \frac{I(box, centroid)}{U(box, centroid)} \tag{6}$$

$$d(box, centroid) = 1 - IoU(box, centroid) \tag{7}$$

2.5. **Insulator defects detection method based on cascaded structure.** The aerial images have complex backgrounds, and the size of insulators and insulator defects area in the images is quite different. Using ordinary object detection algorithm to detect large and small targets at the same time will have a high false detection rate. In order to solve this problem, a novel detection architecture which can not only locate insulators but also detect their defects is proposed. The proposed model transforms insulator defects detection into two cascaded object detection problems, uses the improved YOLOv3 object detection algorithm to locate the insulator positions, then cuts out the insulator area from the original images according to the coordinate information, and finally takes the cut out images as the input of the object detection algorithm of the second stage to detect the defects on the insulators. This method reduces the possibility that the object to be detected in the image is too small to be detected, and improves the accuracy of insulator defects detection. Figure 6 shows the structure of cascaded convolutional neural network in this paper.

## 3. Experimental Results.

3.1. **Model Trainings.** Data sets used in this paper includes:247 defective Insulator picture and 601 standard insulator images. It includes the location information of insulators and insulator defects. Due to the limited size of training samples, the method in reference [21] is used to expand the data to avoid over fitting in the process of model training.

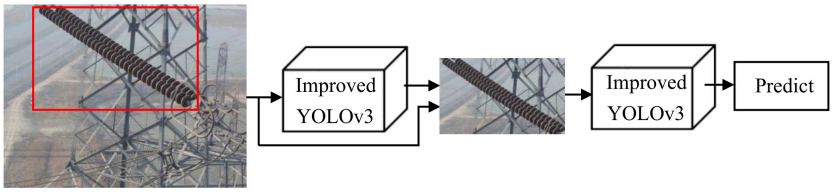


FIGURE 6. Structure of RFB layer

Fig. 7 is an example of the original data set, and Fig. 8 is the image datasets after data expansion.

FIGURE 7. Dataset examples

The experiment is carried out under the operating system of Ubuntu 16.04. Using the deep learning framework pytorch. The CPU used in this experiment is Intel Core i5-7500, graphics card is NVIDIA GTX 1080Ti, memory size is 16GB.

In this paper, 80 % of the above data sets are randomly selected as the training set, and the remaining 20% as the test set. After repeated experiments, it is concluded that the model trained by using the model training parameters in Table 1 has the best detection effect. Fig. 9 shows the network convergence curve when the parameter is used for training.

FIGURE 8. Data augmentation

TABLE 1. Description of training parameters

| Parameter Name | Value |
|---|---|
| Batch Size | 16 |
| Learning Rate | 0.001 |
| weight decay | 0.0005 |
| Optimizer | Adam |
| Momentum | 0.9 |
| MatchThreshold | 0.5 |
| NMS | 0.3 |

3.2. **Experimental evaluation criteria.** In multi class object detection, each class can draw a curve according to recall and precision. AP is the area under the curve, and map is the average value of multi class AP. In object detection, map can well express the accuracy of the algorithm. The calculation of precision and recall rate is as follows:

$$Precision = \frac{T_P}{T_P + F_P} \tag{8}$$

$$Recall = \frac{T_P}{T_P + F_N} \tag{9}$$

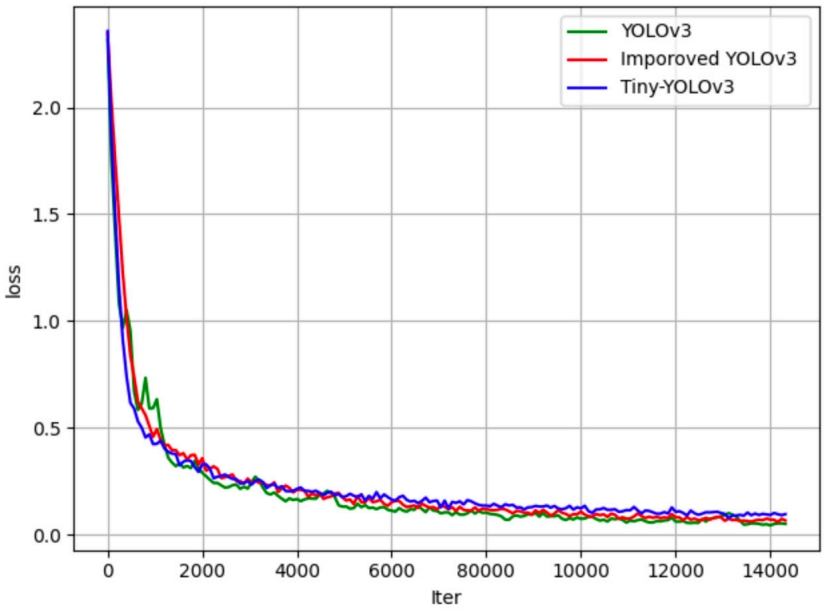


FIGURE 9. Loss curve

$T_P$ is the number of positive samples detected as positive, and $F_P$ is the number of negative samples detected as positive. $F_N$ is the number of Yin and Yang samples tested as negative. Figure 10 shows the precision recall curve in this paper.

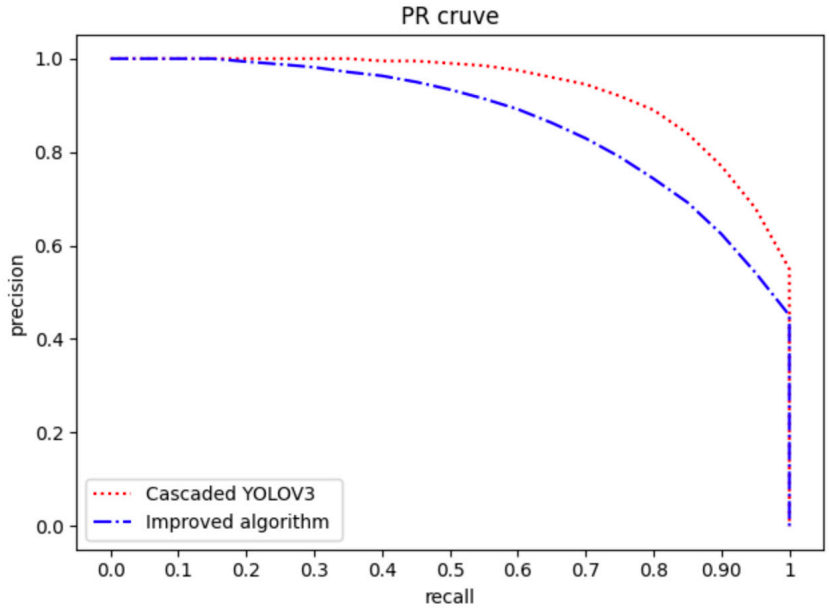


FIGURE 10. Precision-recall curve

Besides detection accuracy, another important performance index of object detection algorithm is speed. The real-time detection is very important for some application scenarios. Frames per second (FPS), the number of pictures that can be processed per second, is a common index to evaluate the speed.

3.3. **Experimental result.** Table 2 lists the model size of YOLOv3, Tiny-YOLOv3 and the improved YOLOv3 proposed in this paper, the average accuracy on the VOC2007

data set and the amount of calculation required to process a picture. It can be seen that compared with Tiny-YOLOv3, the improved network model in this paper has advantages in all aspects. Compared with YOLOv3, although the accuracy has declined, it has greater advantages in terms of model size and calculation volume. Therefore, it is suitable for deployment on mobile devices.

TABLE 2. Performance comparison of different object detection algorithms

| Name | Size(MB) | mAP | GFLOPs |
|---|---|---|---|
| YOLOv3 | 246.5 | 79.2 | 65.7 |
| Tiny-YOLOv3 | 34.7 | 58.4 | 5.52 |
| ImprovedYOLOv3 | 25.8 | 65.1 | 4.72 |

Fig. 11 is a detection effect diagram using single-stage YOLOv3, cascade YOLOv3 and the improved algorithm proposed in this paper. Among them, (a) is the result diagram of normal insulator defects detection, (b) is the result diagram of defective insulator defects detection under simple background, and (c) is the result diagram of defective insulator defects detection under complex background.

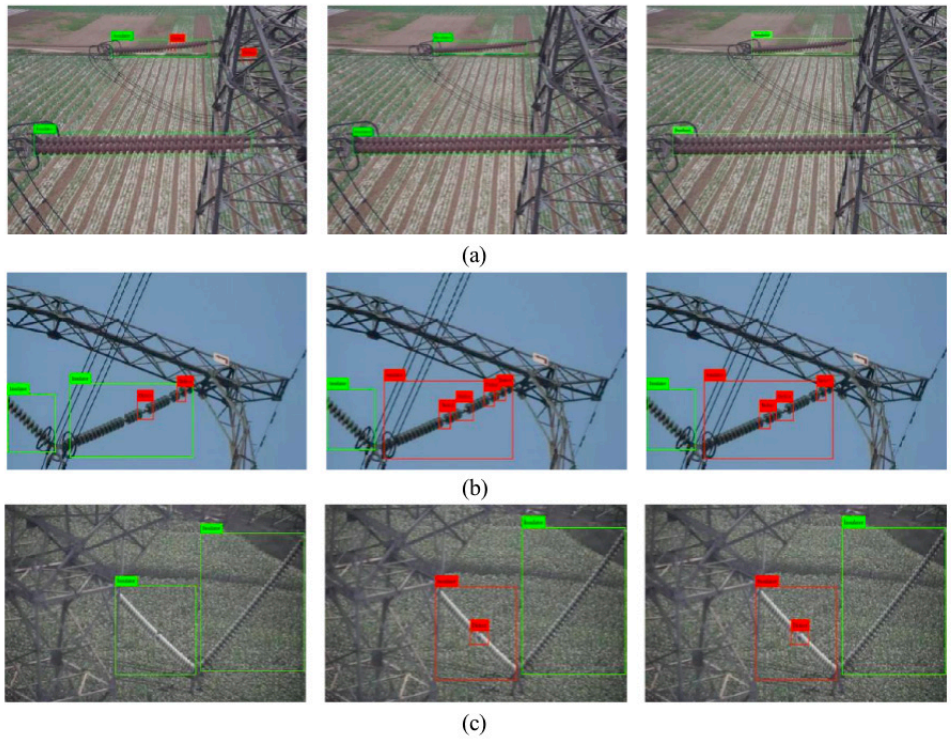


FIGURE 11. Insulator defects detection results

Table 3 lists the detection results of insulator defects detection using different algorithms. It can be seen that the method provided in this paper can ensure the detection speed while maintaining high accuracy. The improved algorithm has significant advantages in detection speed than the original algorithm. Compared with the single-stage detection algorithm, the cascaded algorithm has significantly improved the accuracy of detecting insulator defects. Because the proportion of insulator defects in the whole picture is very small, it is difficult to detect. Therefore, detecting the insulators in the picture first, cutting them, and then using the detected insulators as input to detect the defects

in the insulators can effectively avoid this problem. While maintaining high accuracy, the algorithm realizes real-time detection and model lightweight, and is suitable for embedded edge computing of UAV.

TABLE 3. comparison of insulator defects detection algorithms

| Name | FPS | Insulator (AP@0.5) | Defect (AP@0.5) | mAP (0.5) | Insulator (AP@0.75) | Defect (AP@0.75) | mAP (0.75) |
|---|---|---|---|---|---|---|---|
| Faster R-CNN | 5.6 | 92.05% | 62.33% | 77.19% | 85.72% | 50.66% | 68.20% |
| SSD | 15.6 | 90.69% | 62.57% | 76.63% | 86.78% | 55.98% | 71.38% |
| YOLOv3 | 29.2 | 89.86% | 61.45% | 75.66% | 84.28% | 52.64% | 68.46% |
| Improved YOLOv3 | 126.5 | 86.06% | 64.56% | 75.31% | 82.74% | 56.94% | 69.84% |
| Cascaded Faster R-CNN | 2.4 | 91.94% | 88.38% | 90.16% | 84.49% | 80.56% | 82.56% |
| Cascaded SSD | 4.0 | 90.87% | 87.98% | 89.3% | 86.73% | 83.50% | 85.12% |
| Cascaded YOLOv3 | 8.3 | 90.03% | 86.68% | 88.355% | 85.62% | 81.45% | 83.56% |
| Cascaded Improved YOLOv3 | 59.4 | 86.25% | 87.80% | 87.02% | 82.57% | 79.62% | 81.10% |

4. **Conclusion.** This paper presents an insulator defect detection method based on improved YOLOv3 algorithm. The improved YOLOv3 algorithm uses a new shallow feature extraction network, and uses a deep separable reel to replace the traditional reel operation, which effectively improves the detection speed of the network. The use of the anti-residual structure prevents the gradient from disappearing during the training process, and the detection accuracy is ensured by increasing the RFB layer and predicting on the feature maps of two scales of 13×13 and 26×26. Finally,defect detection of insulators is realized by cascading two improved YOLOv3 algorithms. Experimental results show that compared with other algorithms, this algorithm has the advantage of fast detection speed while maintaining a higher accuracy rate, and can detect insulators in aerial images in real time.

## REFERENCES

[1] 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*, 2021, https://doi.org/10.1007/s12652-020-02740-2.

[2] C. M. Chen, L. Chen, Y. Huang, S. Kumar, M. T. Wu, Lightweight authentication protocol in edge-based smart grid environment, *EURASIP Journal on Wireless Communications and Networking*, 2021, https://doi.org/10.1186/s13638-021-01930-6.

[3] Y. Zhai, D. Wang, M. Zhang, J. Wang, F. Guo, Fault detection of insulator based on saliency and adaptive morphology, *Multimedia Tools and Applications*, vol. 76, no. 9, pp. 12051-12064, 2017.

[4] Y. J. Wang, P. P. Cao, X. S. Wang, 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.

[5] Y. Z. Yun, J. Lou, T. Q. Zhu, H. Wang, AC/DCHybrid System Containing Flexible HVDC Fault Simulation Analysis, *Journal of Northeast Electric Power University*, vol. 37, no. 2, pp. 19-23, 2017.

[6] Z. N. Lv, A survey of common faults analysis and detection methods for transmission line, *Automation & Instrumentation*, vol. 243, no. 1, pp. 167-170, 2017.

[7] L. H. Zhao, Q. Gao, D. H. Li, X. Xiao, Multiple insulators extraction method based on complexinfrared images, *Laser Journal*, vol. 42, no. 5, pp. 62-67, 2021.

[8] G. W. Shao, Z. Liu, J. Fu, J. Y. Tan, Y. Chen, L. W. Zhou, Research Progress in Unmanned Aerial Vehicle Inspection Technology on Overhead Transmission Lines, *High Voltage Engineering*, vol. 46, no. 1, pp. 14-22, 2020.

[9] X. Y. Wang, B. Han, D.D Li, J. Luo, H. Sheng, J. Zhang, Vision based insulator defect detection method, *Computer Engineering and Design*, vol. 40, no. 12, pp. 3582-3587, 2019.

[10] T. Guo, H. Yang, L. Shi, P. Shen, Y. Yang, X. W. Liu, D. Y. Li, T. Z. Zhu, Self-Explosion Defect Identification of Insulator Based on Faster Rcnn, *Insulators and Surge Arresters*, 2019, https://doi.org/10.16188/j.isa.1003-8337.2019.03.031.

[11] H. B. Zheng, J. H. Li, Y. Liu, Y. H. Cui, Y. Ping, Infrared Object Detection Model for Power Equipment Based on Improved YOLOv3, *Transactions of China Electrotechnical Society*, vol. 36, no. 7, pp. 1389-1398, 2020.

[12] Z. Zhao, N. Liu, L. Wang, Localization of multiple insulators by orientation angle detection and binary shape prior knowledge, *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 22, no. 6, pp. 3421-3428, 2016.

[13] S. Ren, K. He , R. Girshick, J. Sun, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2017.

[14] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You Only Look Once: Unified, real-time object detection, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779-788, 2016.

[15] J. Redmon, A. Farhadi, YOLO9000: better, faster, stronger, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7263-7271, 2017.

[16] J. Redmon, A. Farhadi, YOLOv3:An incremental improvement, *arXiv e-prints*, 2018, https://arxiv.org/abs/1804.02767v1.

[17] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, A. Rabinovich, Going deeper with convolutions, *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1-9, 2015.

[18] K. He, X. Y. Zhang, S. Q. Ren, J. Sun. Deep residual learning for image recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, 2016.

[19] A. Howard, M. Sandler, G. Chu, L. C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, Searching for MobileNetV3, *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 1314-1324, 2019.

[20] Y. Zhu, J. Mu, H. Pu, B. Shu, FRFB: Integrate Receptive Field Block Into Feature Fusion Net for Single Shot Multibox Detector, *2018 14th International Conference on Semantics, Knowledge and Grids (SKG)*, pp. 173-180, 2018.

[21] X. Tao, D. Zhang, Z. Wang, X. Liu, H. Zhang, 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.