

# Methods of transmission channel insulator detection based on deep learning model

Liping Lu<sup>a,\*</sup>, Lin Ji<sup>a,b</sup>, Jiangyun Yu<sup>a</sup>, Xinghe Qu<sup>a,b</sup>, Zhimin Yin<sup>a</sup>

<sup>a</sup>State Grid Huzhou Electric Power Supply Company, Huzhou, 313000, China

<sup>b</sup>Zhejiang Tailun Electric Power Group Co., Ltd., Huzhou, 313000, China

## ABSTRACT

Using artificial intelligence technology to realize intelligent detection and monitoring of transmission lines is one of the key technologies in the construction of new power system. This paper introduces the target detection of transmission channel based on deep learning method. The deep learning method has the advantages of little influence of super parameters on the results, strong feature extraction ability and strong anti-interference ability. This paper mainly introduces the network framework, including YOLO, SSD, R-FCN, fast R-CNN, etc., and analyses the advantages and disadvantages of various methods. At the same time, the transmission channel insulator detection is verified for various network architectures, and the experiments show the applicability of various types to transmission channel target detection.

**Keywords:** Insulator, deep learning, defect detection, transmission channel, computer vision

## 1. Introduction

In recent years, machine vision technology based on deep learning<sup>1,2</sup> has been greatly developed, and the corresponding image target detection algorithm has also achieved good performance. After Alexnet achieved extraordinary performance in the image recognition competition in 2012, the deep learning algorithm with convolutional neural network (CNN) as the core has become the main research direction of image classification and target detection<sup>3</sup>. Target detection algorithms based on deep learning can be divided into “double order method” and “single order method”. “Double order method” has high accuracy, and representative algorithms include Fast R-CNN<sup>4</sup> and R-FCN<sup>5</sup> (R-FCN algorithm is improved on the basis of Faster R-CNN); The “single order method” is fast, and the representative algorithms are YOLO<sup>6,7</sup> and SSD<sup>8</sup>. We will focus on the application of these four algorithms to transmission line target detection.

## 2. Four Mainstream Algorithms

### 2.1. Transmission line target detection based on YOLOv3

YOLO algorithm directly regresses the target location and target category in the output layer to realize end-to-end training and detection, which is different from the original two-order target detection method based on region recommendation. YOLOv3 adds multi-scale prediction, which makes the network have stronger detection ability for targets with large size change range, and have higher detection speed and recognition accuracy. For any machine patrol picture, first adjust the size to a unified size, and divide the picture into  $S \times S$  regions, and then extract picture features through multiple convolution layers. For each region, if the center of the target falls within this region, the region is responsible for predicting the target, and then adjust the center position of the detection target and the size of the prediction frame through regression. The output result of model prediction is  $S \times S \times (B \times 5 + C)$ , precisely  $S \times S$  areas. Each area outputs the position of  $B$  defect prediction box with different sizes and  $C$  category information, and each prediction box has 4 coordinate values and 1 confidence value. Finally, the model uses maximum suppression to remove duplicate boxes, and finally predicts the type and location of detection targets actually contained in the picture, as shown in Figure 1.

Zhou and Fang<sup>9</sup> and others proposed a rapid defect detection method of transmission line based on YOLOv3 in their paper. Firstly, the YOLOv3 target detection model is established, and the a priori size of the target candidate region is obtained by cluster analysis of the training sample database; By adjusting the loss function to speed up the training

\*13511220161@163.com

process of the model and adjust the learning direction of the model. Finally, through the establishment and training of transmission line defect sample database, transmission line defect detection (defect category is 20 categories) is realized. In this paper, the author adjusts two parameters compared with the original YOLOv3 model to make the model more suitable for transmission line defect detection. Firstly, for the input image, the detection area is divided according to three scales, there are  $19 \times 19$ ,  $38 \times 38$  and  $76 \times 76$  detection areas respectively (using dark-53 as the basic network and forming a characteristic map of three scales according to the size of the input picture), k-cluster the defect size to obtain 9 cluster centers, and the values of these 9 cluster centers are used as the sizes of 9 a priori boxes. Apply larger priori boxes (including  $512 \times 337$ ,  $261 \times 272$ ,  $137 \times 144$ ) on the smallest  $19 \times 19$  feature graph, suitable for detecting large objects. Apply medium priori boxes (including  $39 \times 281$ ,  $127 \times 77$ ,  $63 \times 66$ ) on the medium  $38 \times 38$  feature graph, suitable for detecting medium-sized objects. Apply smaller priori boxes (including  $47 \times 33$ ,  $26 \times 20$ ,  $7 \times 9$ ) on the larger  $76 \times 76$  feature graph, suitable for detecting small objects. Therefore, the first improvement is to obtain 9 priori frames suitable for defect size through k-clustering. Then, in the design of loss function, the author considers that the area with target is far smaller than the area without target, and set the weight  $\lambda$  of no target position to 0.1. There are 20 kinds of defects in the model design, but there are only 4 coordinate parameters. Directly adding up the coordinate loss, category loss and confidence loss will lead to too little influence of coordinate parameters on the loss function. Authors set the coordinate loss weight to 5. The second improvement is the adjustment of the weight of the loss function. Finally, through experimental analysis, the overall accuracy is improved by 4% compared with the original YOLOv3 model.

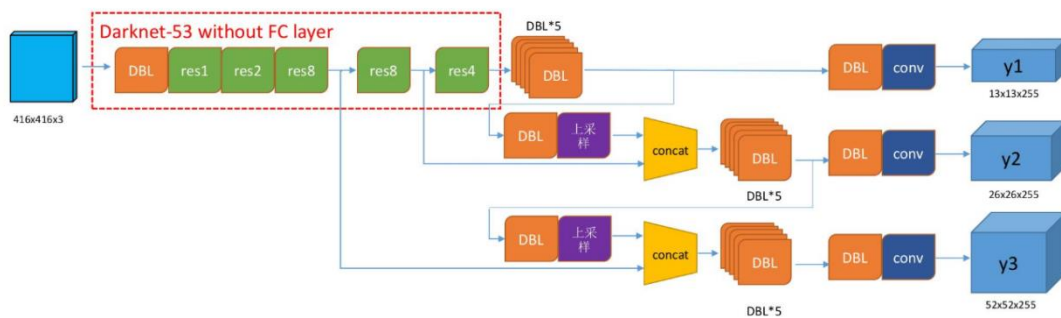


Figure 1. YOLOv3 structure.

Weng, Cheng and Zheng<sup>10</sup> and others proposed a YOLOv3 algorithm model based on Res2Net residual<sup>11</sup> module in the paper on the detection method of key parts of high voltage transmission line based on improved YOLOv3. The algorithm mainly makes two improvements to the original algorithm. Firstly, the algorithm introduces Res2Net<sup>12</sup> residual module to replace the original residual module (a group of cascaded  $3 \times 3$  convolution layer, divide the previously input feature map into  $s$  groups. Each group is a convolution channel, and the output of the previous channel will be fused before convolution, so that richer feature information can be extracted. Theoretically, the more channels  $s$ , the finer the extracted feature granularity, but when  $s=4$ , the overall performance of the network is the best, Therefore, in this paper,  $s$  is set to 4), which uses fewer modules to obtain higher feature extraction efficiency, effectively improves the fine-grained detection ability of the network, and solves the problems of diversified transmission inspection objectives and changeable angles. Then it draws lessons from the transfer learning scheme of knowledge distillation (first train the yov3 model, then set the trained yov3 model as the teacher model, and the improved algorithm YOLO res as the student network. Through knowledge distillation, YOLO res learns on the trained yov3 model, and the obtained model is the pre training model used in network training in this paper), The fusion of shallow information is strengthened (the detection scale of the network is relatively small, so the obtained features are mostly advanced features. The author hopes to obtain the location information features from the shallow network. After testing, the characteristic maps output by the Res2Nst residual module of  $208 \times 208$  scale are respectively transmitted to the convolution layers of the three scales in the second half of the network by means of maximum pooling, which fully integrates the shallow features and advanced features), which improves the accuracy and reduces the missed detection rate compared with the original network. Finally, the algorithm is applied to the self-built aerial photo data set. The experimental results show that the improved network based on yolov3 can realize the synchronous and rapid positioning of key patrol components. The average accuracy of the model constructed by this method is 90.9% and the running speed is 42frame/s, especially in small target detection. YOLO res is 1.8 percentage points higher than yolov3's map index, The efficiency index FPS increased by 7.7%.

## 2.2. Transmission line target detection based on SSD

Single shot multibox detector (SSD) is a classical model in the deep learning target detection model. The SSD model outputs discrete multi-scale and multi-scale extraction frames, adopts multi-scale feature maps and carries out classification and regression tasks on multiple feature maps at the same time, and realizes the comprehensive improvement of speed and accuracy. SSD is a single target detection and recognition method, which is similar to YOLO. It only needs to process the picture once to achieve multi-target detection and recognition, and the accuracy is higher than YOLO. The depth network feature maps with different scales are used to predict the target at each location, which ensures that the accuracy of detection and recognition can be compared with the method based on regional recommendation. Even if the image is a low-resolution image, SSD algorithm can still achieve high-precision detection and recognition. SSD uses the full convolution VGGNet et as the backbone network to directly act on the feature map to predict multi-target categories and bounding boxes. In order to balance the accuracy and real-time of detection, SSD adopts the single point multi frame detection method instead of the candidate region method used in Faster R-CNN. Even for images with relatively low-resolution input, it has good detection effect. The network structure is shown in Figure 2.

Li, Zheng and Zhao<sup>13</sup> and others proposed a method for insulator defect detection<sup>14</sup> on transmission line based on improved SSD algorithm in the research on insulator defect detection method on transmission line based on SSD algorithm. Firstly, the data set is extended, and preprocessing such as clipping and rotation are used to improve the generalization ability of network detection and realize image annotation. Secondly, ResNet<sup>15</sup> is used to replace the original VGGNet<sup>16</sup> in SSD structure, so as to enhance the feature extraction ability of the network. Finally, the overlapping area of insulator position and insulator defect position detected by convolution neural network is calculated to determine the insulator defect position. The author mainly improves the SSD algorithm in two aspects. Firstly, the original SSD model is based on VGG-16 as the basic network for detection, and then the convolution kernel is used to predict the category and coordinate offset of a series of bounding boxes on the feature maps of different scales. On the basis of retaining the original target detection basic network, the author replaces the original feature extraction network VGG-16 with ResNet-101, To a certain extent, it deepens the overall structure of the network, enhances the ability of network feature extraction, and improves the average accuracy compared with the original algorithm. On the other hand, the calculation of IOU in an overlapping area is introduced at the end. Because the false detection rate of defective insulators is high and many of the false detection parts are not on the insulator, the author can calculate the IOU of the two according to the physical relationship between the defective part and the insulator. If the IOU is greater than 0.9, it is considered that the detected position of the defective.

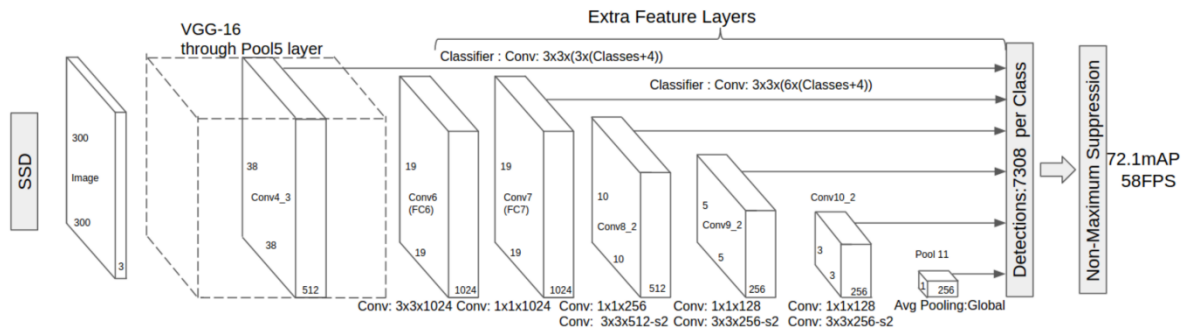


Figure 2. SSD structure.

## 2.3. Transmission line target detection based on faster R-CNN

Faster R-CNN model consists of feature extraction, regional candidate network and Fast R-CNN network. Firstly, the model scales the input image; A set of basic convolution layer, excitation layer and pooling layer are used to extract the feature map of the input image, and the feature map is used for the subsequent RPN layer and full connection layer; RPN network is used to generate candidate regions. Firstly, several anchor frames are generated, cut and filtered, and the foreground or background is judged through softmax, that is, whether it is a target or not; At the same time, the other branch, bounding box region, modifies the coordinate information of the anchor box to form a more accurate candidate region; After the candidate regions are generated by RPN network and the overlapping distribution of candidate regions is suppressed by NMS algorithm, the candidate regions are classified by Fast R-CNN network to determine whether they

are detection targets, and the boundary coordinates of the measured targets are determined by multi task loss frame regression algorithm. It is an end-to-end network<sup>17</sup>. The network structure is shown in Figure 3.

Yin, Meng and Fan and others<sup>18</sup> proposed a typical visual defect detection system for substation equipment based on Faster R-CNN, which integrates edge calculation<sup>19-21</sup> and improved Faster R-CNN. The author first makes two improvements to Faster R-CNN, namely NMS<sup>22</sup> algorithm and bounding box generation mechanism. Firstly, the obtained data set is flipped, translated and other enhanced data operations, and the obtained data set is expanded to reduce the occurrence of over fitting and other problems. The original NMS algorithm will lead to the elimination of targets with low confidence, resulting in missed detection of targets. The author optimizes the NMS algorithm. When the IOU value is  $\geq$  NMS threshold (the original NMS algorithm will set the boundary box score of IOU greater than the threshold to 0, that is, remove these boundary boxes, and the improved algorithm will reduce the boundary box scores greater than the threshold rather than all set to 0, plus setting a weight to slow down the attenuation of scores), the confidence will be reduced. Slowing down the attenuation of confidence and insulator is effective and the detection of the optimized defective part is realized. The experimental results show that the accuracy of the improved algorithm is improved by 4.77%, and through the calculation of the position relationship between the insulator and the defective part, the false detection rate of the defective insulator is reduced and the detection accuracy is improved.

Weight reduction and finally achieve the purpose of improving positioning accuracy. The original RPN candidate box generation area will take the center of the sliding window as the benchmark, and each candidate area with 9 different sizes and proportions will be considered. The author analyses the existing substation image data set, sets five different aspect ratios and four sizes, and considers 20 candidate boxes with different sizes and proportions for the center of each sliding window, so as to further improve the recognition rate of typical defects of small target substation equipment. Finally, the improved Faster R-CNN model is stored in the edge device<sup>23-24</sup> (edge computing means that the data to be processed is directly completed in the edge device, and the processed results are transmitted to the cloud server. The corresponding is cloud computing. Cloud computing<sup>25</sup> means that all data are transmitted to the cloud server for processing to obtain the results. Edge computing can ensure that the number of data transmission, request response time and data security are greatly reduced. The edge device directly detects the defects of the collected video images of substation equipment, and sends the detected defect results to the cloud server. Compared with the traditional model of directly transmitting the image information to the cloud server for detection, the data transmission volume of the proposed detection system is reduced by more than 90%. The experimental results show that the proposed system has typical visual defects in substation equipment The trap has the characteristics of high detection accuracy and strong real-time. Compared with the original Faster R-CNN model, the map value is increased by 6.9%, which can effectively improve the work efficiency and intelligent level of substation patrol inspection.

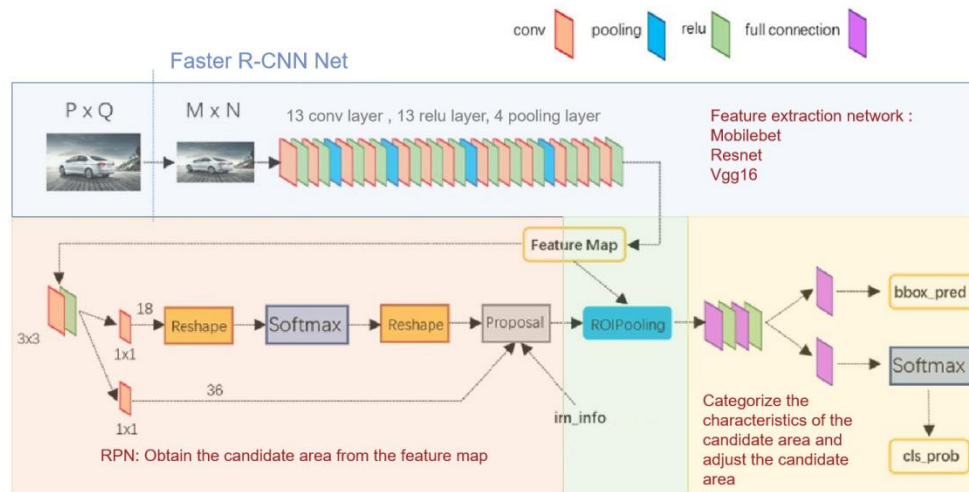


Figure 3. Faster R-CNN structure.

## 2.4. Transmission channel target detection based on R-FCN

R-FCN algorithm is an improvement of Faster R-CNN algorithm. Compared with Faster R-CNN algorithm, R-FCN algorithm has two improvements: first, replace ROI Pooling layer with “location sensitive” area pooling layer to resist the loss of target location information caused by pooling operation; Secondly, before the regional pooling layer, the score of the probability size and relative position of the target belonging to each category is generated through the convolution neural network. The two scores are mapped to the same size through the ROI Pooling layer. By averaging the mapped scores, the probability and relative position of each region belonging to different categories are directly obtained. In this algorithm, the “location sensitive” region pooling layer does not need to pool the features of each dimension in the whole target region (just calculate the maximum value) and only selects a small region for operation. Therefore, the extracted features are more refined and retain more location information; In addition, it is not necessary to calculate the full connection layer for the features extracted from each region, so the amount of calculation is smaller than that of Fast R-CNN. The network structure diagram is shown in Figure 4.

Zhao, Cui and Qi<sup>26</sup> studied the insulator detection method in the aerial line inspection image based on the improved R-FCN. In this paper, the insulator target in the aerial line inspection image is partially blocked, and the regional full convolution network (R-FCN) model is used to detect it. The generated region based features are used as the input of the countermeasure network. For the target features, ASDN<sup>27</sup> attempts to generate a mask to indicate which parts of the features to be discarded, so as to generate “more difficult” and more diversified samples for the detector. The mask layer is introduced to form a mask pair feature map through the residual convolution layer. Partial occlusion is performed, and the obtained results are against the network results before the ASDN layer is introduced.

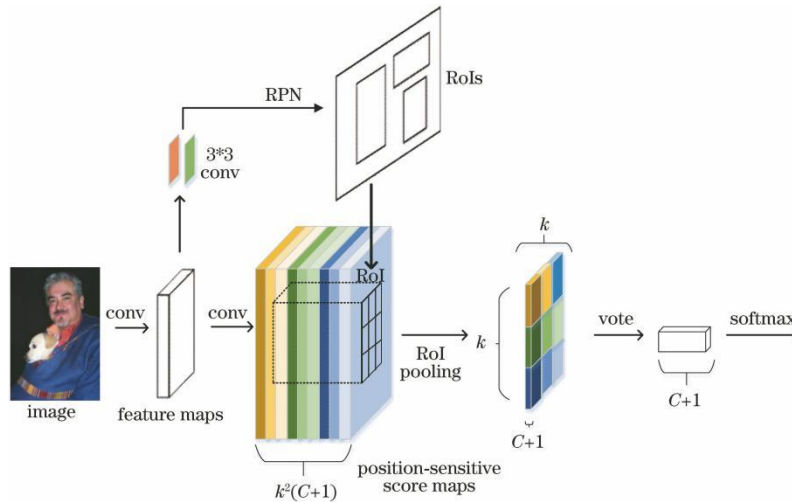


Figure 4. R-FCN structure.

## 3. EXPERIMENTAL RESULTS

In the experiment, 599 insulator pictures obtained from UAV patrol inspection and manual patrol inspection are used as the data set, 80% of them are randomly selected as the training set, and the remaining 20% of them are used as the verification set. The training set is input into YOLOv3, SSD, Faster R-CNN and R-FCN for training. When the loss no longer decreases, the loss weight at that time is selected as our training model. Finally, the pictures of the verification set are input into the four models obtained from training for testing, and finally the output results of different networks are obtained, as shown in Figure 5. The results obtained are analyzed and compared, and a table is made. From Table 1, we can see the performance comparison between networks. It can be seen from the table that Faster R-CNN has the highest mAP, but its speed is relatively slow because it is a two-order algorithm. The fastest network is SSD, but its map value is a little worse than Faster R-CNN.

Table 1. Experimental comparison diagram.

Model	Backbone	Input size	mAP-50/%	FPS	Weight size
YOLOv3	DarkNet53	416x416	93.55	13.1	235M
SSD	VGG16	300x300	93.85	20.3	90.6M
R-FCN	ResNet-101	224x224	93.63	3.4	351M
Faster R-CNN	ResNet50	600x600	96.04	3.3	315M

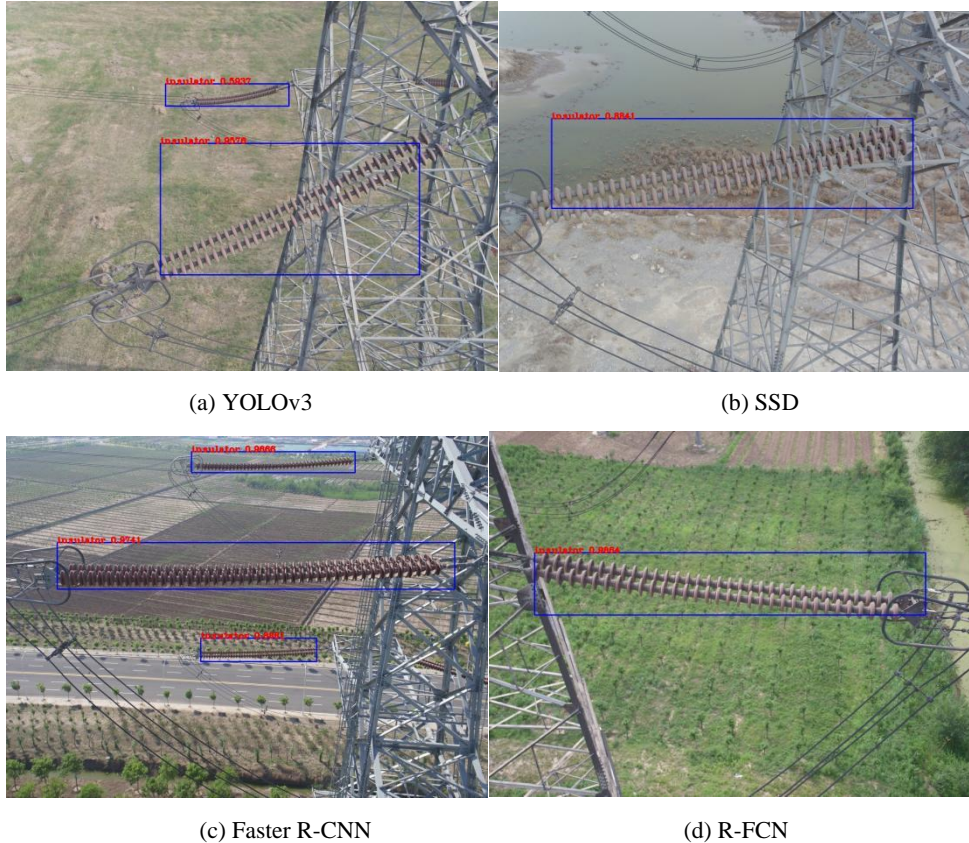


Figure 5. Output results of four networks.

#### 4. CONCLUSION

Nowadays, the application of deep learning method to transmission line target detection has been widely promoted, and has been applied to practice to realize the intellectualization of detection. With the development of computer vision technology, some new deep learning algorithms and technologies will be gradually applied to the field of power industry, so as to expand the content of transmission line target detection and improve the level of power informatization and intelligence.

#### REFERENCES

- [1] Szegedy, C., Toshev, A. and Erhan, D., "Deep neural networks for object detection," Advances in Neural Information Processing Systems, 2553-2561 (2013).

- [2] Krizhevsky, A. and Hinton, G., "Learning multiple layers of features from tiny images," Technical Report, University of Toronto, 1-60 (2009).
- [3] Simonyan, K. and Zisserman, A., "Very deep convolutional networks for large-scale image recognition," Computer Science, arXiv preprint arXiv:1409.1556 (2014).
- [4] Ren, S., He, K., Girshick, R. and Sun, J., "Faster R-CNN: Towards real-time object detection with region proposal networks," Transactions on Pattern Analysis and Machine Intelligence, 39(6), 1137-1149 (2017).
- [5] Dai, J., Li, Y., He, K. and Sun, J., "R-FCN: Object detection via region-based fully convolutional networks," Conf. and Work. on Neural Information Processing Systems (NIPS), 379-387 (2016).
- [6] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., "You only look once: Unified, real-time object detection," Computer Vision and Pattern Recognition (CVPR), 779-788 (2016).
- [7] Redmon, J. and Farhadi, A., "YOLOv3: An incremental improvement," arXiv preprint arXiv:1804.02767, (2018).
- [8] Liu, W., Anguelov, D., Erhan, D., et al., SSD: Single shot MultiBox detector Computer Vision 21-37 (2016).
- [9] Zhou, F. R., Fang, M., Ma, Y. T. and Pan, H., "Fast detection method of transmission line defects based on YoloV3," Yunnan Electric Power Technology, 48(4), 6 (2020). (in Chinese)
- [10] Weng, Z., Cheng, X. and Zheng, Z. Q., "Detection method of key components of high voltage transmission line based on improved YoloV3," Computer Application, 40(S2), 184-187 (2020).
- [11] He, K., Zhang, X., Ren, S. and Sun, J., "Deep residual learning for image recognition," CVPR2016: Proc. of the 2016 IEEE Conf. on Computer Vision and Pattern Recognition, IEEE Computer Society, Washington, DC, 770-778 (2016).
- [12] Gao, S., Cheng, M., Zhao, K., Zhang, X. Y. and Torr, P. H. S., "Res2Net: A new multi-scale backbone architecture," Transactions on Pattern Analysis and Machine Intelligence, 43(2), 652-662 (2019).
- [13] Li, W. X., Zheng, W. L., Wang, N. and Zhou, H. H., "Research on insulator defect detection method on transmission line based on SSD algorithm," Instrument Users, 26(8), 1-4 (2019). (in Chinese)
- [14] Szegedy, C., Toshev, A. and Erhan, D., "Deep neural networks for objects detection," Advances in Neural Information Processing Systems, 2553-2561 (2013).
- [15] Szegedy, C., Ioffe, S., Vanhoucke, V. and Alemi, A. A., "Inception-v4, inception-ResNet and the impact of residual connections on learning," Thirty-First AAAI Conference on Artificial Intelligence, 4278-4284 (2017).
- [16] Wang, L., "Places205-VGGNet models for scene recognition," Computer Science, arXiv preprint arXiv:1508.01667 (2015).
- [17] Tao, W., Wu, D. J., Coates, A. and Ng, A. Y., "End-to-end text recognition with convolutional neural networks," 2012 21st Inter. Conf. on Pattern Recognition (ICPR), 3304-3308 (2012).
- [18] Yin, Z. H., Meng, R., Fan, X. D., Li, B. and Zhao, Z. B., "Typical visual defect detection system for substation equipment based on edge computing and improved fast R-CNN," China Science Paper, 16(3), 6 (2021). (in Chinese)
- [19] Shi, W., Jie, C., Quan, Z., Li, Y. and Xu, L., "Edge computing: Vision and challenges," IEEE Internet of Things Journal, 3(5), 637-646 (2016).
- [20] Shi, W. S., Zhang, X. Z., Wang, Y. F., et al., "Edge computing: Current situation and prospect," Computer Research and Development, 56(01), 73-93 (2019). (in Chinese)
- [21] Li, B., Jia, B. C., Chen, S., et al., "Application prospect of edge computing in power supply and demand," China Power, (11), 154-162 (2018). (in Chinese)
- [22] Ren, S., He, K., Girshick, R. and Sun, J., "Faster R-CNN: Towards real-time object detection with region proposal networks," IEEE Transactions on Pattern Analysis & Machine Intelligence, 39(6), 1137-1149 (2017).
- [23] Lu, Y. Q., Sun, C. Y., Cao, H. W., et al., "Foreign object detection method of transmission equipment based on edge calculation and deep learning," China Power, 53(6), 27-33 (2020). (in Chinese)
- [24] Liu, X. L., He, S. H., Pang, Z. J., et al., "Research on intelligent infrared inspection device based on edge computing," Electrical Age, 458(11), 47-49 (2019). (in Chinese)
- [25] Ma, F. Q., Wang, B., Dong, X. Z., et al., "Power vision edge intelligence: Power depth vision acceleration technology driven by edge computing," Power Grid Technology, 44, 17-26 (2020). (in Chinese)
- [26] Zhao, Z. B., Cui, Y. P., Qi, Y. C., Kong, Y. and Zhang, K., "Insulator detection method in aerial line inspection image based on improved R-FCN," Computer Science, 1-15 (2019).
- [27] Wang, X., Shrivastava, A. and Gupta, A., "A-Fast-RCNN: Hard positive generation via adversary for object detection," Computer Vision and Pattern Recognition (CVPR), 3039-3048 (2017).