



*Research article*

## **An improved Faster R-CNN for defect recognition of key components of transmission line**

Hongxia Ni <sup>1</sup>, Minzhen Wang <sup>1,\*</sup> and Liying Zhao <sup>2</sup>

<sup>1</sup> School of Electrical Engineering and Information Technology, Changchun Institute of Technology, Jilin 130012, China

<sup>2</sup> School of Computer Technology and Engineering, Changchun Institute of Technology, Jilin 130012, China

\* **Correspondence:** Email: 147371426@qq.com; Tel: +086-139-4416-6477; Fax: +86-431-8057-8188.

**Abstract:** In a national power grid system, it is necessary to keep transmission lines secure. Detection and identification must be regularly performed for transmission tower components. In this paper, we propose a defect recognition method for key components of transmission lines based on deep learning. First, based on the characteristics of the transmission line image, the defect images are preprocessed, and the defect dataset is created. Then, based on the TensorFlow platform and the traditional Faster R-CNN based on the R-CNN model, the concept-ResNet-v2 network is used as the basic feature extraction network to improve the network structure adjustment and parameter optimization. Through feature extraction, target location, and target classification of aerial transmission line defect images, a target detection model is obtained. The model improves the feature extraction on transmission line targets and small target component defects. The experimental results show that the proposed method can effectively identify the defects of key components of the transmission lines with a high accuracy of 98.65%.

**Keywords:** convolutional neural networks; defect identification; Faster R-CNN; Inception-ResNet-v2; target detection

---

### **1. Introduction**

Electricity has gradually become an indispensable energy source in our lives and production. Scholars worldwide have carried out considerable research work on fault detection and key

component identification of transmission lines. Sun Jin used the slope model to achieve edge matching and Canny to perform edge detection [1]. Zhang et al. used a variety of transformation methods, such as expansion corrosion, to address insulator treatment, which improved the accuracy of insulator troubleshooting [2]. Chen Xiaojuan, used the Hough transform to detect specific parts and then performed further positioning work according to the appearance characteristics of shock hammers [3]. Murthy and other scholars used wavelet transform to achieve insulator feature extraction and obtain a fault recognition method based on the Markov model [4]. In the field of insulator fault identification, Zheng Tao et al. used the pulse-coupled network [5]. Under the guidance of the concept of region segmentation, the grey value of each block was compared with a preset threshold. Based on the AdaBoost algorithm, Jin Lijun and other researchers created shock hammer image recognition [6]. Haar-like features can extract and classify the relevant features of dampers so that high accuracy matching can be achieved in the detection process. Cui and other experts improved the pulse-coupled neural network and used the improved algorithm to identify insulator components [7]. Oberweger and other scholars used cluster analysis to distinguish insulator target images from the perspective of statistics, which can be effectively separated from the target, and the ellipse descriptor can be used to effectively observe whether the insulator has defects [8].

In 2015, Zhang Fengyu et al. used contour matching and other means to carry out insulator component image recognition [9]. The first step corrected and preprocessed the input image, converted the original RGB color to HSV color successfully, and used the H component as the matching space. Next, the vertical and horizontal directions of the filter can be used to effectively determine whether the insulator string drops. This kind of method can build a model to complete the identification work, but the practicability is not strong, and the operation is difficult. In 2017, Miao Xiangpeng et al. studied and discussed the differences between the current deep learning model processing method and previous machine learning [10] and found that previous machine learning methods mainly used feature training and classifiers to effectively recognize images, but the current deep learning model is very advanced in concept and can achieve a higher accuracy recognition rate than traditional methods. Based on the SSD model, Wang Sen and others optimized the identification system of shock hammers [11]. With the help of the random Hough transform, this kind of method is only suitable for image recognition in a specific background. Although in this experiment, the author can achieve 97% accuracy in massive data, the defect is also very clear and the scope of the application is not wide. Wang Wanguo [12] and other scholars have attempted to identify various power components through a deep learning framework. This paper uses a variety of current mainstream deep learning frameworks and carries out actual measurements, which can realize the accurate identification of transmission line components and obtain a relatively satisfactory recognition speed. In 2019, Hou Chunping [13] and others used the most representative deep learning target classification algorithms, such as AlexNet, VGG16 and Faster R-CNN, to train the classifier and detector respectively, and cascaded the classifier and detector to form a cascade network for insulator target detection and recognition. In 2020, Wei Shuangfeng et al. Studied the automatic identification and defect detection methods of transmission lines and insulators based on aerial images in view of the shortcomings in the defect detection of transmission lines and their typical components insulators. [14]. Xiong Xiaoping et al. Studied the transmission line defect recognition model based on Faster R-CNN [15]. Firstly, the establishment process of image database for transmission line defect recognition is described in detail. Then, by analyzing and comparing the performance of three pretraining front-end networks and the influence of multiple parameters on the

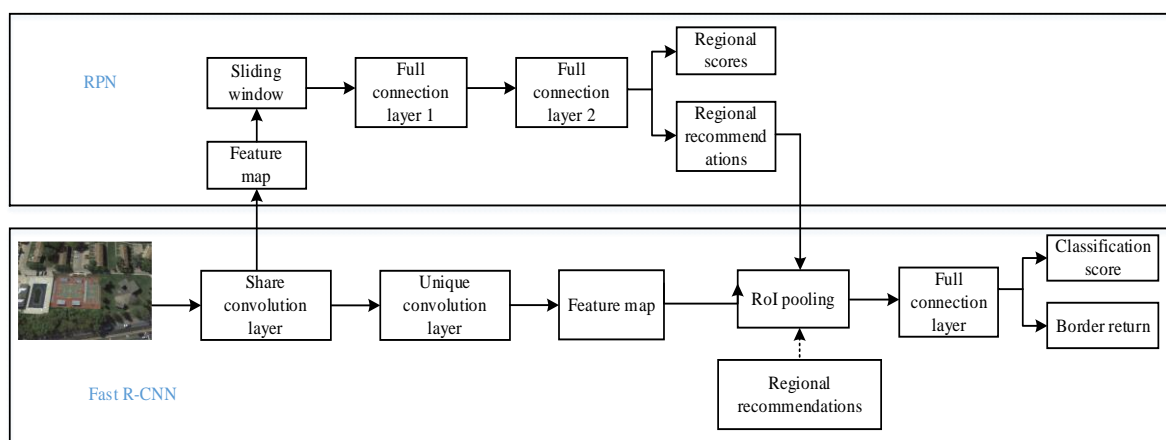
accuracy of network model recognition, the optimal network model for transmission line defect recognition based on Faster R-CNN is obtained.

In this paper, according to the characteristics of defect images of key components of transmission lines, the image is preprocessed, and the defect dataset of the key components of the transmission line is made. Under the TensorFlow platform, the excellent classification network of the Perception-ResNet-v2 network is optimized. As the basic network of the two-stage detector fast RCNN, it is used for defect recognition of key components of transmission lines, and it is the most effective in the experimental results compared with different algorithms; this method can effectively improve the accuracy and reliability of defect identification of key components in transmission line inspection.

## 2. Faster-RCNN

### 2.1. Basic principles

Fast R-CNN [16] needs to search a large number of fault node feature data in the application, which leads to a long process and low fault recognition efficiency. To strengthen the fault detection ability of the method and further improve the detection efficiency, reference [17] improved the method, and proposed a new Fast R-CNN detection method, which used the regional proposal network (RPN) to identify the defects, thus greatly improving the efficiency of defect identification. The combination of RPN and Fast R-CNN forms a Fast R-CNN detection link, as shown in Figure 1.



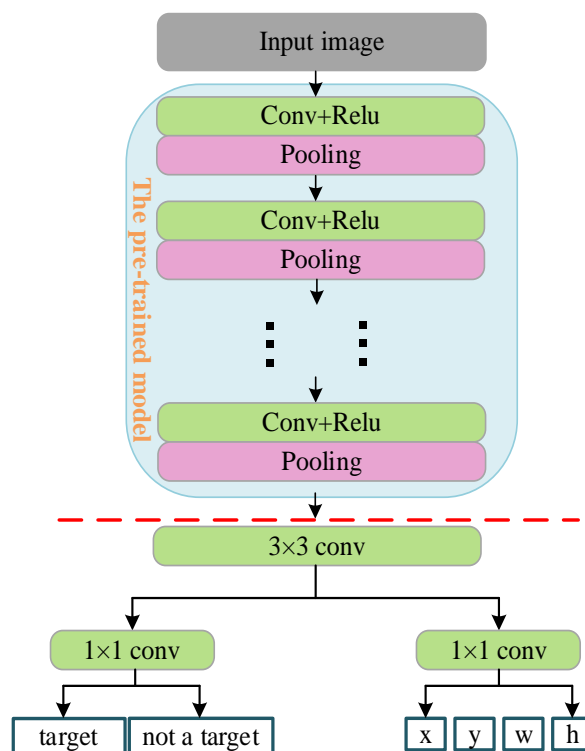
**Figure1.** Faster R-CNN structure.

As seen in Figure 1, Faster R-CNN first inputs the image to be detected into a convolutional neural network (such as VGG-16). After arriving at the shared convolution layer, the feature map is divided into two directions. One direction is used as the input of RPN to generate candidate regions, and the other continues to arrive at the unique convolution layer to generate a higher-dimensional feature map. In the practical application of RPN, the fault area range of insulators is obtained by inputting the characteristic graph of the fault node in the sliding window area, and by evaluating the score of fault elements in the area, the non-maximum value is inserted to suppress the fault parameters, and approximately 300 fault candidate areas are screened out and sent to the ROI pooling layer. The ROI pooling layer not only contains the RPN candidate regions but also integrates

the volume base high-dimensional feature data, which can extract the key features from the target region. Finally, the fault node data are obtained through border regression and region classification to realize the analysis and identification of fault range and fault type.

## 2.2. Regional recommendation network optimization

The essence of the RPN network is a sliding fault window, which contains a large number of fault pixels. Each pixel has an independent scale, making it a multi-scale anchor. Because the target of self-explosion defect of the transmission line identified in this paper is small, three different scales (162, 322, 642) and different aspect ratios (1:1, 1:2, 2:1) are adopted in this paper. There are nine kinds of anchors in total. The coordinate position of the original image is obtained by translating the anchor coordinates. The output nine anchor classes have specific position coordinates, and each coordinate has an independent foreground and background. The window regression layer can output nine anchors at all positions, corresponding to the parameters  $(x, y, w, h)$  of the moving and zooming windows, where  $x, y, w, h$  represent the center point coordinates and width height of the window respectively, as shown in Figure 2.



**Figure 2.** RPN network structure.

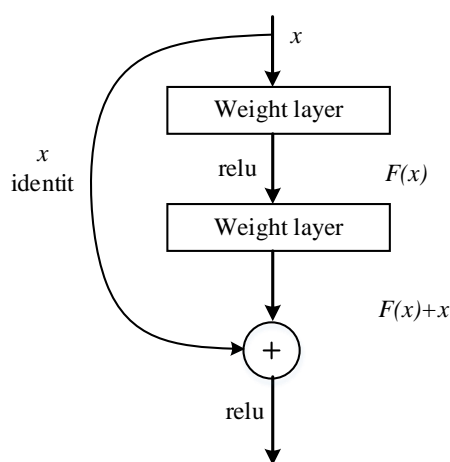
To train the RPN network better, this paper assigns class labels {foreground, background} to each anchor and uses the non-maximum suppression method to obtain the prediction box. For positive labels, the anchor corresponds to the maximum value of the overlapped IOU and the anchors corresponding to the IOU greater than 0.7 of any labelled border; for negative labels, the IOU of any labelled border is less than 0.3. The corresponding anchor cannot be used as a positive or negative label, and the anchor in the convenient position of the image can be ignored. Finally, RPN and Faster

R-CNN share features, and input ROI pooling classifies and regresses the prediction box. According to the particularity of pointer defect detection, RPN is used to generate a smaller scale preselector box to ensure convergence direction and accuracy.

### 3. Inception-ResNet-V2 architecture and optimization

#### 3.1. Inception-ResNet-V2 structure

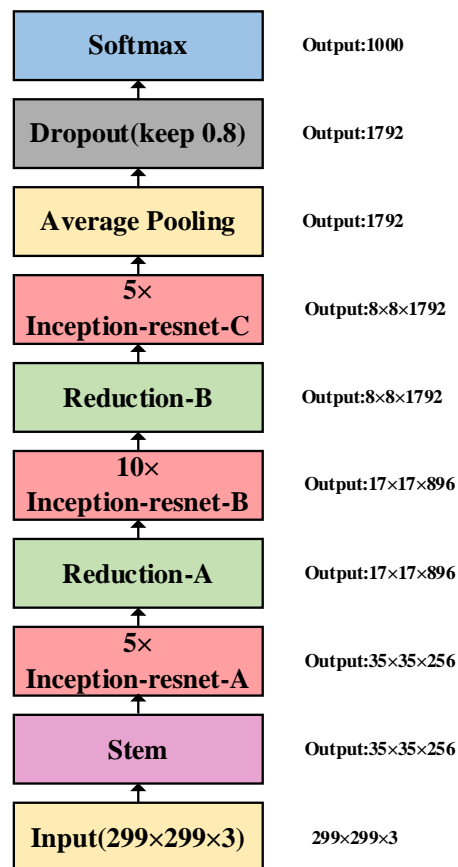
In this paper, Inception-ResNet-V2 is introduced into the research of defect identification of key components of transmission lines, and the basic network structure of defect identification is constructed. Inception-ResNet [18] integrates the Inception network structure and ResNet network structure, which is essentially a combination of the two network structures. The network structure model of the Inception [19] series is a model proposed by Google. It does not limit its use of specific convolution cores but uses all convolution cores of different sizes simultaneously and then stitches the resulting matrix together. To further improve the computational efficiency, it is necessary to introduce a  $1 \times 1$  convolution kernel before the  $3 \times 3$  junction to simplify the number of convolution kernel channels and optimize the convolution dimension. After multiple convolution processing, the stem structure generated by Inception V4 is also processed twice by pooling. Pooling takes the convolution pooling parallel structure proposed by Inception-v3 [20] as the core, which can greatly reduce redundant convolution parts. ResNet [21] is a 152-layer residual neural network structure proposed by Kaiming he in the ImageNet competition. For more complex problems, a deeper neural network is needed, but deeper networks cause problems such as overfitting, gradient dispersion, and gradient explosion. The residual neural network introduces a shortcut structure, which adds the input layer transfer and convolution results, and the problem of gradient dispersion caused by too deep of a neural network is alleviated. The residual element is shown in Figure 3.



**Figure 3.** Residual unit.

Inception-ResNet-V2 is not only the optimization and improvement of InceptionV4, but also the introduction of ResNet. The architecture of the model is divided into input blocks, Inception-Resnet-C, and Reduction-B blocks. The network structure composed of Inception-Resnet-A and Inception-Resnet-B is improved in both accuracy and operation efficiency.

The operation efficiency is almost the same as that of Inception V4, and the operation precision is higher, and the operation result is closer to reality. The overall structure is shown in Figure 4.



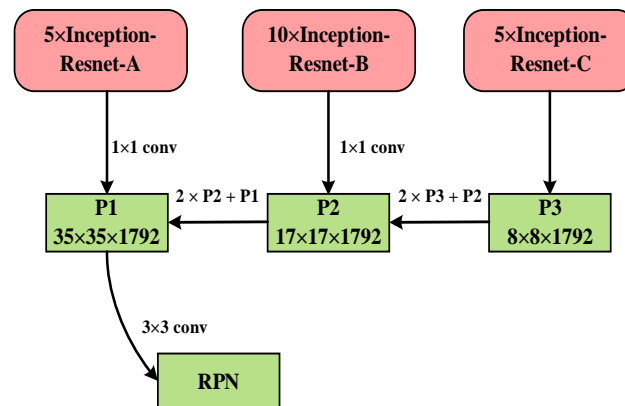
**Figure 4.** Inception-ResNet-V2 overall structure.

### 3.2. Multi-scale feature fusion

The conventional structure of Faster R-CNN is decomposed into multiple levels by pooling layer processing, and each level has an independent scale, which makes it achieve higher recognition accuracy. However, there is a problem: because the different depths in different layers affect the size of the receptive field in each layer, the extracted image feature information also has different characteristics. In the low-level structure, it is easy to achieve more integrity, so it has more favorable information for target positioning. However, in the high-level structure, it is more inclined to obtain the information of target classification, so there is a large gap between the semantic information of high-level and low-level. Therefore, to obtain the features that can be obtained by using as many layers as possible, they should be fused to achieve the best target detection results.

The commonly used scale fusion learning structure mainly consists of four types of graph scales: image pyramid, multi-level feature map, multi-scale feature fusion, and feature pyramid. Among them, the image scale under the image pyramid structure mainly realizes the expression of image scale information through the size of the image resolution, which has great time and space costs. For the Faster R-CNN network, the training time may be too long or the memory may be insufficient. SSD integrates a variety of scale feature fusion methods and does not need the support of image feature sampling. It can select representative image features from different network layers and

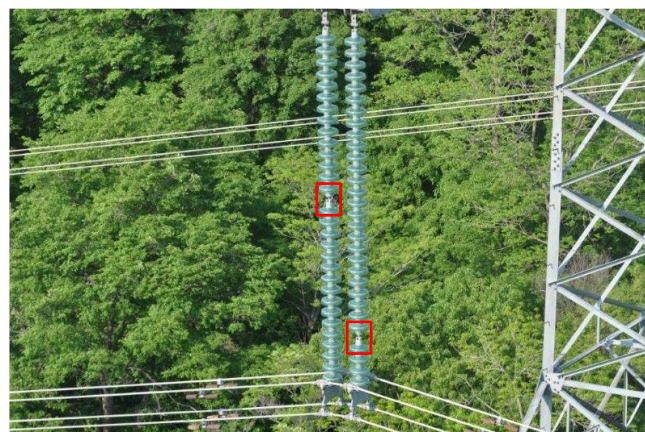
predict the change trend and direction of image features and needs very few image feature points to be calculated. According to the feature pyramid principal, this paper optimizes the structure of the traditional Inception-ResNet network, as shown in Figure 5, and outputs characteristic graphs of Inception-ResNet-A and Inception-ResNet-B are dimensionalized to 1,792 of Inception-ResNet-C by a  $1 \times 1$  convolution kernel. Inception-ResNet-C processes the image, outputs the image with feature samples, and samples the image features more than twice. Then, Inception-ResNet-B integrates and optimizes the image features. Finally, by fusing with P1, the mixed features are obtained based on  $3 \times 3$  convolution processing and are introduced into the RPN layer.



**Figure 5.** Inception-ResNet feature fusion map.

### 3.3. Convolution layer optimization

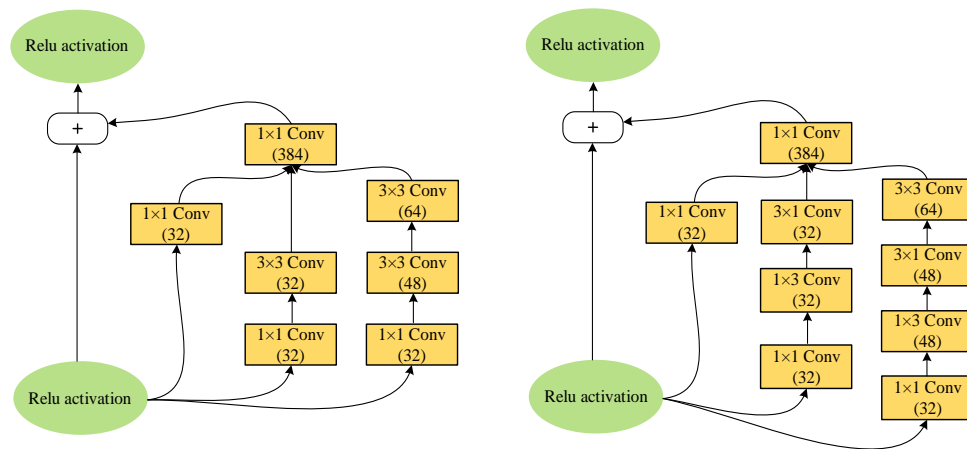
By comparing and observing the defects of the key components in the image, it can be found that the size of the target defect image is small, as shown in Figure 6. When the convolution layer adopts the minimum convolution kernel, it is difficult to identify the defect features.



**Figure 6.** Defect image.

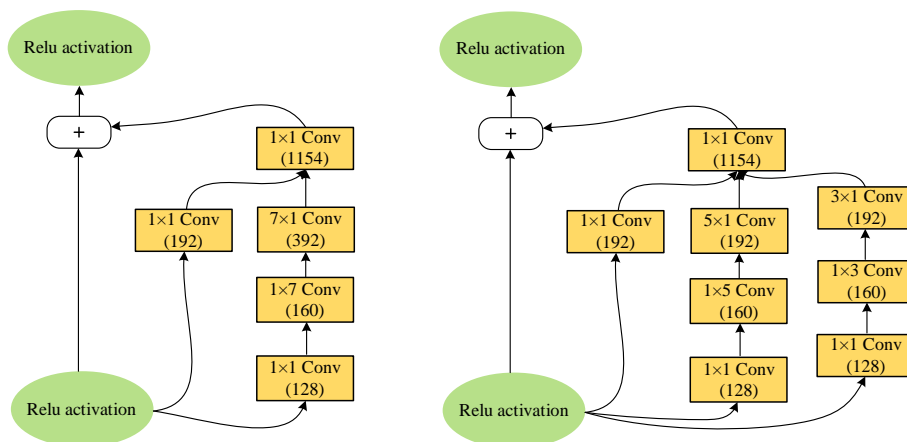
To improve the recognition accuracy, the following structure optimization is proposed for the traditional Inception-ResNet-V2 network. In the Inception-ResNet-A module, the convolution kernel size is  $3 \times 3$ , which can be refined into two independent convolution kernels of  $1 \times 3$  and  $3 \times 1$ , thus simplifying the calculation process of defect characteristic parameters. As shown in Figure 7,

the  $3 \times 3$  convolution in branch B and the first  $3 \times 3$  convolution in branch C are divided into two  $1 \times 3$  convolutions, while the second  $3 \times 3$  convolution in branch C remains unchanged so that the network depth does not change as much as possible.



**Figure 7.** Inception-ResNet-A module optimization.

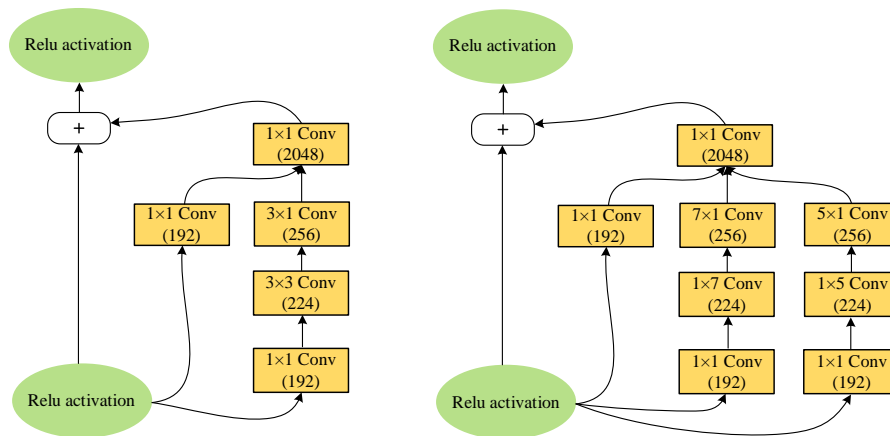
For Perception-ResNet-B, the original  $1 \times 7$  convolution core is changed to convolution core  $1 \times 5$ , as shown in Figure 8, and a new branch is added. The  $1 \times 3$  convolution structure is adopted, that is, the width of the convolution layer is increased, and multiple receptive fields are used to capture different levels of feature information. Finally, they are summarized together.



**Figure 8.** Inception-ResNet-B module optimization.

Since a large number of  $3 \times 3$  and  $5 \times 5$  convolution kernels are used in Inception-ResNet-A and Inception-ResNet-B, we hope that they can extract higher-order abstract features in Inception-ResNet-C. therefore, the  $3 \times 3$  kernel in the original Inception-ResNet-C is replaced by  $7 \times 7$  kernels, and a  $5 \times 5$  branch is added to aggregate large-area feature information in the image, as shown in Figure 9.





**Figure 9.** Inception-ResNet-C module optimization.

## 4. Data set and model training

### 4.1. Image preprocessing

Due to the UAV in the process of outdoor high-altitude shooting, it is easily affected by wind, rain, light and other factors, resulting in image blur, uneven light and shade, noise pollution and other problems, so that the image quality is not high. If these data are directly used as the input of subsequent image recognition deep learning, it will often affect the experiment because it cannot provide enough features. The image obtained in this paper is affected by the weather, and there are different light and dark; in the process of outdoor image shooting, there must be noise pollution and light interference. To reduce the negative impact on the image quality, the aerial image is preprocessed.

In the process of histogram equalization, the most critical step is to calculate the probability of the grey value in different image regions and then make the only image. Combined with the image stretching processing, the number of pixels in a certain range is the same to achieve the effect of peak contrast enhancement and valley bottom contrast weakening. The method of color image histogram equalization is the same as that of grey images. Both of them distinguish and recognize the image regions of red, green, and blue. The process of histogram equalization is as follows:

If the grey level of the original image at  $(x, y)$  is  $f$  and the grey level after equalization is  $g$ , the process can be expressed as mapping the grey level  $f$  at  $(x, y)$  to  $g$ . The mathematical method can be defined as formula (1).

$$g(x, y) = T[f(x, y)] \quad (1)$$

The mapping function  $T(r)$  needs to meet the following two conditions at the same time:

- (1) The value range of  $T$  increases monotonically on the grey level  $0 \sim L-1$ ;
- (2) The value range of  $T$  is between 0 and  $L-1$ , where  $L = 256$ .

There is an important function in image processing, as shown in formula (2), which meets the above conditions. This function is called CDF (cumulative distribution function), and its core function is to describe the different variables in a specific image area distribution probability.

$$s = T(r) = (L - 1) \int_0^r p_r(w)dw \quad (2)$$

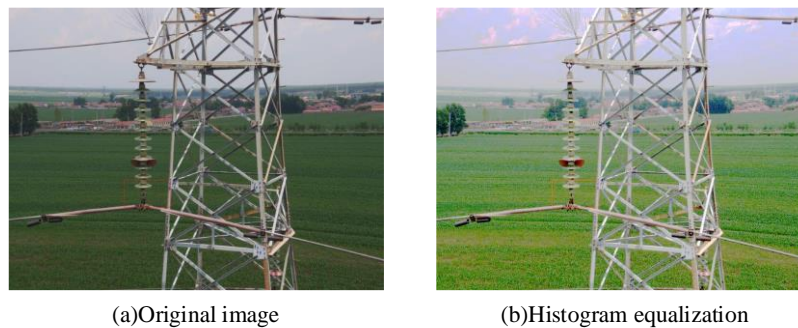
Since the image pixel distribution is a discrete function, formula (2) can be written as formula (3):

$$s = T(r) = (L - 1) \sum_{i=0}^r p_r(i) \quad (3)$$

where  $p_r(i)$  represents the probability of the grey level occurring in the image. Assuming that the total number of pixels in the image is  $n$  and the number of pixels in each grey level in the histogram is  $h(i)$ , it is finally expressed as formula (4).

$$s = T(r) = (L - 1) \sum_{i=0}^r \frac{h(i)}{n} \quad (4)$$

The image is processed by the above method, and the result after the histogram equalization is shown in Figure 10.



**Figure 10.** Equalization result graph.

It can be seen in Figure 10 that compared with the original image, the characteristics of the insulator are more obvious in the picture after image enhancement, and the background is more prominent. By adjusting the image brightness, some small and unimportant parts in the background can be removed, thereby weakening the background.

#### 4.2. Dataset preparation

The original data of the transmission line image database were captured by UAVs in the Jilin Province Jiangjing line, Jiahe line 2, Bei'an Beixi line, Dunming line, Pingyan line, Tianqiao line, Tiantao line, Qingyong line, and Rebai line. It includes 1,321 insulator images, 650 hammer images, and 161 bird nest images. There are 512 insulator defects and 235 shock hammer defects. The size of the original image of the transmission line is  $6,000 \times 4,000$  pixels. Due to the limited ability of GPU training, it is difficult to train a large number of samples, so the original image is compressed to  $800 \times 600$ . The characteristics of defects in key components of transmission lines are shown in Figure 11.

In this paper, open source LabelMe software is used to mark all the defect images collected, as shown in Figure 12. The peripheral box of the defect location is marked, the coordinate position of the peripheral box is recorded, and the category label is assigned. In the image, 80% of the region is defined as the image training set, and 20% of the region is defined as the test set. After statistical analysis, a total of 1,008 key component defects were marked.



**Figure 11.** Insulator defect image.



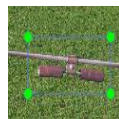


**Figure 12.** Annotated dataset.

#### 4.3. Description of defect identification

This paper mainly identifies the defects of insulators, bird nests, and anti-vibration hammers in transmission lines. The specific description is shown in Table 1. Self-explosion defects of insulators, corrosion of anti-vibration hammers, and bird's nest defects are identified. The number of each defect in the dataset is also shown in Table 1.

**Table 1.** Description of defect identification.

Defect category	Defect mark name	Defect diagram	Number of defects
Insulator defect	Defective insulator		592
Bird's nest defect	Bird nest		161
Defects of shock hammer	Vibration damper		255
Total			1008

#### 4.4. Training process and parameter setting

The training process of the improved Faster R-CNN model in this paper follows the following four steps. (1) The improved pre-training Inception-ResNet-V2 model is used to initialize the convolution network. First, the RPN network is trained independently, and then the `first_stage_nms_iou_threshold` is obtained. The initial parameter of the threshold is set to 0.7, and `first_stage_max-proposals` are obtained. When max proposals are set to 300, 300, candidate boxes are obtained, and the Faster R-CNN network learning rate is set to 0.0002. (2) The pre-trained Inception-ResNet-V2 model is used to initialize the convolution network, and the candidate region in step (1) is used as the input to initialize the Faster R-CNN network. The two networks do not share the convolution layer. (3) The convolution layer parameters updated by using the loss function in step (2) are used to redefine an RPN network, which is a convolution network, the learning rate of the R-CNN network is set to 0.0002, and only the RPN network is updated. (4) The candidate region generated in step (3) is taken as the input, and the updated convolution layer in step (2) is sent to the Faster R-CNN network to form a unified network. After extracting the internal features of the network, the RPN is used to generate the candidate region, and the Faster R-CNN network identifies and locates the features in the candidate region.

Through the above process, the RPN network and Faster R-CNN network share a convolution layer, and other parameter settings in the training process are shown in Table 2.

**Table 2.** Other parameter settings.

Name	Method/Parameter
Weight	Normal distribution initialization (0~1)
learning_relu	SGD
learning_rate	0.0002
learning_momentum	0.9
learning_policy	step
batch_size	1
num_steps	200000

## 5. Experimental results and analysis

### 5.1. Experimental environment

The experiment is based on one of the most popular deep learning frameworks, TensorFlow, which is a convenient, efficient, extensible, lightweight software that can run on multiple platforms. The software is easy to install, easy to learn, developed and maintained by Google and has a strong community. It can refer to and solve a series of problems in project implementation. It supports multiple GPUs to accelerate the training process. Its built-in TensorBoard visualization software can track the topology and performance of the network. Therefore, model problems can be identified as soon as possible and solved in time. The software and hardware configuration used in the experiment is shown in Table 3.

**Table 3.** Experimental environment parameters.

Processor	RAM	Display card	Operating system	framework	programming language
IntelCore i5-4210@2.4GHz	16G	GeForce GTX 980	Win 10 64bt	TensorFlow 1.4-gpu	Python2.7

### 5.2. Evaluation index

In the process of comparing the application effects of all target detection technologies, it is not only necessary to compare the detection results with the naked eye, but also to select some objective evaluation indicators as the evaluation basis for the validity of the detection results. In most cases, AP (average precision index), mAP (mean average precision), and recall rate (recall rate) can be used as control indicators.

Assuming that there are only two types of test results, positive samples, and negative samples, the test results can be divided into four cases, as shown in Table 4.

**Table 4.** Confusion matrix.

	Prediction positive	Prediction negative
Actually positive	TP	FN
Actually negative	FP	TN

Accuracy indicates how many of the results predicted to be positive samples are true samples, and it indicates the performance of the detection model on the ability to determine the target category, its calculation formula is shown in formula (5).

$$P = \frac{TP}{TP+FP} \quad (5)$$

The recall rate indicates how many of the positive samples in the sample are correctly predicted. It measures the ability to detect the model on the ability to recognize the target. Its calculation formula is shown in formula (6).

$$R = \frac{TP}{TP+FN} \quad (6)$$

When there are multiple targets in the detection task, the average accuracy of each target is often different. Therefore, to comprehensively consider the detection effect, the average accuracy of all targets can be used as the basis for solving the average value. The area under the PR curve is defined as AP, as shown in formula (7). The mean value of AP under all categories is calculated to obtain the map value, as shown in formula (8).

$$AP_i = \int_0^1 p(r) dr \quad (7)$$

$$mAP = \frac{1}{k} \sum_{i=1}^k AP_i \quad (8)$$

### 5.3. Comparative analysis of experimental results

In this paper, the Faster R-CNN training method and Inception-ResNet-V2 pre-training method are used to process the image in the process of image feature recognition of key components of

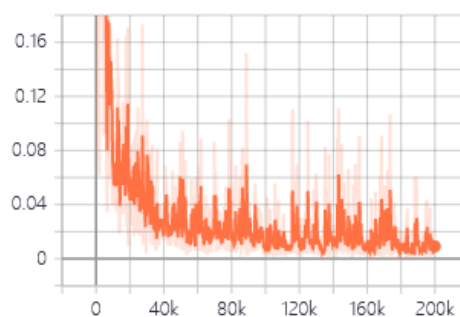
transmission lines, and the best performance of Faster R-CNN is obtained. The results show that the accuracy recognition rate of this method is 98.65%, and the recall rate is 96.45%. The effective test time of a single image is 676 ms, which shows that this method has high accuracy and efficiency and has strong practicability.

To prove the accuracy and effectiveness of the algorithm, the combination of different feature extraction networks and detection models is used in the transmission line defect recognition image, and the comparison results of accuracy and recall rate under different combination models are obtained. Through the result analysis, as shown in Table 5, in terms of test time, the test speed of the SSD model is faster than that of Faster R-CNN, but the recognition accuracy is not high. The algorithm in this paper is higher than other combined models in accuracy, and the test time is also within the acceptable range, so the combination model of Faster R-CNN and Perception-ResNet-v2 can obtain a better recognition effect.

**Table 5.** Different combinations of model accuracy Compare.

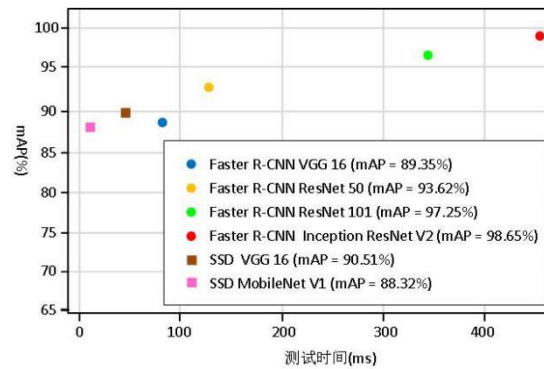
Model name	Pre-trained model	Map /%	Recall /%	test time /ms
	RseNet-101	97.25	95.09	508
Faster	RseNet-50	93.62	91.33	120
R-CNN	VGG-16	89.35	90.76	90
	Inception-ResNet-V2	<b>98.65</b>	<b>96.45</b>	676
SSD	VGG-16	90.51	89.22	53
	MobileNet-V1	88.32	87.82	26

The loss function can reflect the learning quality of the model. By observing the convergence process of the loss curve, the model gradient disappearance and gradient explosion can be understood. TensorBoard is used to observe the loss curve that appears during the training process to determine the training completion. Figure 13 shows the loss curve obtained after 200 000 iterations.



**Figure 13.** Loss curve.

In the early stage of training, the drop speed of the loss value is faster, and when it reaches 0.08, it goes into a fluctuating state. The more iterations, the smaller the loss value. After 10 K iterations, the numerical fluctuation decreases. Although some data have large fluctuations, the impact on the overall convergence effect is very limited. Therefore, the learning effect of the model can be evaluated based on the change characteristics of the loss curve. Figure 14 shows the test time and accuracy of each comparison model in this paper.



**Figure 14.** Accuracy and test time.

Figure 15 shows part of the detection results of this method. The detection results cover the identification results of the data sets of the key components of the transmission line under different backgrounds. The results show that this method can effectively detect the defects of the key components of the transmission line, and has good recognition performance.



**Figure 15.** Recognition results in this article.

## 6. Conclusion

This paper first analyses the traditional Faster R-CNN target recognition method and then improves the traditional method. Considering the characteristics of the image defects of the key components of transmission line that need to be identified, Inception-ResNet-V2 is used as the feature extraction pre-training network of Faster R-CNN, and the network optimization and parameter adjustment of the pre-training network are carried out. The experimental results show that the improved network can effectively improve the operation efficiency of the network, and the recognition accuracy of the small target of transmission line faults also significantly improved, reaching 98.65%, which provides a new idea for improving the power grid inspection efficiency and ensuring normal operation.

## Acknowledgments

This work was supported by the Jilin Provincial Department of science and technology under Grant 20190302036GX.

## Conflict of interest

The authors have no conflict of interest.

## References

1. J. Sun, Research on insulator crack diagnosis based on edge detection, *North China Electric Power University*, 2008.
2. Y. Zhai, Y. Wu, H. Chen, A method of insulator detection from aerial images, *Sen. Tra.*, **7** (2014), 1726–5479.
3. X. Chen, Y. Wu, L. Zhao, OPGW anti-vibration hammer identification based on random Hough transform, *Heilongjiang Electric Power*, **32** (2010), 1–2.
4. V. S. Murthy, S. Gupta, Digital image processing approach using combined wavelet hidden markov model for well-being analysis of insulators, *Iet Image Process.*, **5** (2011), 171–183.
5. T. Zheng, Research on segmentation and positioning of aerial insulator images based on PCNN, *Dalian Maritime University*, 2011.
6. L. Jin, S. Yan, Y. Liu, Anti-vibration hammer identification based on HAAR-like features and cascaded AdaBoost algorithm, *J. Simul.*, **24** (2012), 1806–1809.
7. K. Cui, B. Li, J. Yuan, An improved unit-linking PCNN for segmentation of infrared insulator image, *App. Mat. Inf. Enc.*, **8** (2014), 2997–3004.
8. M. Oberweger, A. Wendel, H. Bischof, Visual recognition and fault detection for power line insulators, *19<sup>th</sup> Computer Vision Winter Workshop*, 2014.
9. F. Zhang, Recognition and research of abnormal inspection map of transmission line based on computer vision, *Jilin University*, 2015.
10. X. Miao, Identification and damage detection of catenary insulators based on image processing, *Southwest Jiaotong University*, 2017.
11. S. Wang, Research on detection algorithm of anti-vibration hammer on transmission line image, *Beijing Jiaotong University*, 2017.
12. W. Wang, B. Tian, Y. Liu, L. Liu, J. Li, Research on recognition of power widgets in UAV inspection images based on RCNN, *J. GEO-INF.*, **19** (2017), 256–263.
13. C. Hou, H. Zhang, W. Zhang, Y. Yang, G. Zhang, Z. Tian, Identification method of transmission line insulator self-explosion defects, *J. Electron. Syst. Autom.*, 2019.
14. S. Wei, S. Huang, W. Cao, X. Xi, P. Wang, Transmission line insulator recognition and defect detection based on aerial image, *Eng. Invest.*, **48** (2020), 39–43.
15. X. Xiong, S. Xu, D. Meng, X. Wei, D. Tu, W. Wu, Research on transmission line defect recognition model based on faster R-CNN, *Autom. Instrument.*, 2020.
16. R. Girshick, Fast R-CNN, *Proceedings of the IEEE international conference on computer vision*, (2015), 1440–1448.
17. S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks, *IEEE T. Pattern Anal.*, **39** (2017), 1137–1149.
18. C. Szegedy, V. Vanhouck, S. Ioffe, J. Shlens, Z. Wojna, Inception-v4, inception-ResNet and the impact of residual connections on learning, arXiv:1602.07261.
19. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Rethinking the inception architecture for computer vision, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2016), 2818–2826.



20. M. J. Bharata, R. Meyur, D. Pal, An on-line geographical information system–based condition monitoring system for 11-kv distribution line insulator, *IEEE Electr. Insul. M.*, **33** (2017), 26–32.
21. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2016.



AIMS Press

©2021 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)