



# Target Detection in Infrared Image of Transmission Line Based on Faster-RCNN

Shifeng Yan, Peipei Chen, Shili Liang<sup>(✉)</sup>, Lei Zhang, and Xiuping Li

School of Physics, Northeast Normal University, Changchun 130024, China  
lsl@nenu.edu.cn

**Abstract.** Compared with the visual image, the infrared image of the transmission line has lost some image characteristics and the image resolution is lower. In this paper, an improved Faster-RCNN method is used to locate the target in the infrared image of the transmission line. We first construct the infrared image data set of the transmission line and extract the image features by comparing different network models; then we increase the scale and candidate frame when generating target candidate regions in the region proposal network according to the small target features of the infrared image data set. The accuracy of the insulator string (AP) is improved by about 8.4%, and the average accuracy (mAP) is improved by about 3%. Experiments show that this method has higher recognition accuracy when detecting infrared image targets with lower resolution.

**Keywords:** Transmission line · Infrared image · Faster-RCNN · Target detection · Feature extraction network · Candidate box

## 1 Introduction

With the widespread application of unmanned aerial vehicle technology, the disadvantages of traditional manual methods have become more and more obvious, and it is becoming more and more convenient to use aircraft to check [10]. Infrared thermal imaging temperature measurement technology is widely used in the fields of power equipment detection and fault identification due to its characteristics such as small influence by electromagnetic fields, high detection efficiency, and safety in temperature measurement. Infrared detection technology is also widely used in the detection of transmission line faults.

The complexity of the transmission line environment makes it difficult to locate and identify the components of the transmission line. Zou [22] proposed a method for identifying and detecting bird's nests on transmission towers that integrates corner points, straight lines, colors, and shapes. Yan et al. [15] proposed an improved Otsu algorithm based on morphological methods to segment transmission line images, and then used a new filtering method to remove tiny noises according to the geometric characteristics of power lines. Zhu et al. [21] used color space conversion, Otsu segmentation algorithm and edge detection methods to mark the connected domain of the insulator string. Zhao et al. [20] used NSCT's gray entropy model to realize the automatic positioning of insulator

strings in complex backgrounds. Liu et al. [8] used the relative position relationship between the power tower and the inspection aircraft as a priori information to roughly locate the power tower and then used machine learning to further locate it. Yetgin et al. [16] and others used a new strategy based on discrete cosine transform to detect power lines in visible light images or infrared images. Tong et al. [12] proposed a segmentation and identification method for insulators based on aerial images, which can effectively identify insulators operating online with high accuracy.

The traditional method of target detection method is more complicated, and the recognized target is relatively single. Deep learning has made great contributions in many applications [1, 19], such as Yue et al. [2, 3, 17, 18] proposed a series of deep learning-based models that effectively recognize human intentions via EEG signals analysis and have achieved brilliant recognition results. Deep learning theory is widely used in the fields of image recognition and target detection. Wang et al. [13] and others constructed a spiking neural network and designed a new infrared image edge detection method using the characteristics of spiking neurons. In 2014, Girshick et al. [4] first proposed the Region-CNN algorithm. First, the selective search algorithm (Selective Search, SS) extracts the target candidate area, then the deep convolutional network extracts the features, and finally the target category and location are output. In the continuous optimization based on Fast-RCNN [5], Faster-RCNN [9] has been proposed successively. Faster-RCNN algorithm uses RPN to complete the extraction of target candidate regions instead of SS algorithm, which realizes end-to-end training and detection. With the development of deep learning, deep learning methods have gradually been applied to the detection of transmission lines. Wang et al. [14] realized the positioning and identification of low-power components through the RCNN algorithm. Lin et al. [6] can maintain high recognition accuracy and speed in detecting images with different resolutions and different position angles by using the improved Faster-RCNN algorithm. Liu et al. [7] used Faster-RCNN algorithm to locate heating faults in infrared images of power transmission based on the image library of heating faults of power transmission. Tao et al. [11] and others proposed a new type of deep convolutional neural network cascade architecture, which can effectively detect insulator defects under various conditions.

Faster-RCNN is not widely used in power equipment, and there are few studies on using Faster-RCNN to detect the fault of power equipment in infrared images. Therefore, this paper adopts the improved Faster-RCNN algorithm, first establishes a database through infrared images of infrared video clips obtained by infrared cameras, and then realizes the identification and positioning of transmission line components.

## 2 Target Detection Algorithm Based on Infrared Image

### 2.1 Transmission Line Target Detection Algorithm

The combination of RPN and Fast-RCNN can be regarded as Faster-RCNN. RPN realizes the selection of target candidate regions, and Fast-RCNN realizes the classification and positioning of candidate regions. The two neural networks share the convolutional layer, and the feature maps of the convolutional layer are paired. By adjusting the two neural networks, the target detection and positioning of the infrared image are finally realized. The algorithm flow chart is shown as in Fig. 1.

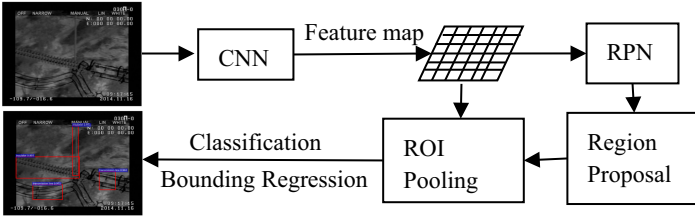


Fig. 1. Faster-RCNN algorithm flowchart

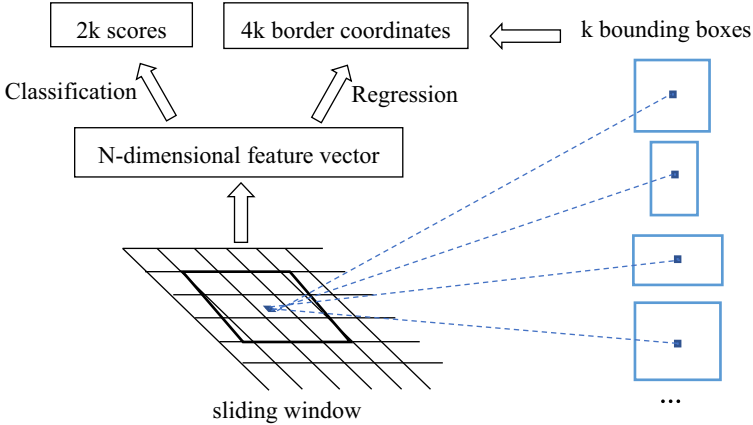


Fig. 2. RPN network

The RPN network searches for all target candidate regions on the feature map. The RPN network is shown in Fig. 2. For the feature map, the RPN generates 9 anchor boxes according to different proportions and different areas, and finally the sum of the anchor boxes generated by all pixels is  $k$ . The feature map is convolved with the  $3 \times 3$  convolution kernel and sliding convolution, and then the category judgment and position determination are performed through two fully connected layers. The regression layer outputs the coordinates of  $k$  boxes; the classification layer outputs the probability of whether there is a target in the anchor box.

When the RPN network generates the candidate area, it uses the non-maximum suppression algorithm to remove the redundant candidate frame, and finally outputs the target candidate area with a higher score as the suggested area to the Fast-RCNN network; Finally, the fully connected layer calculates the classification score and boundary regression to realize the positioning in the recognition.

## 2.2 Faster-RCNN Structure Parameter Selection Optimization

### Feature Extraction Network Model

The VGG16 is practical and performs well in the field of image recognition. They designed a residual module to train a deeper network. The residual module establishes a direct connection between input and output. MobileNet consists of two independent modules,  $3 \times 3$  depthwise Conv ( $3 \times 3$  Depthwise Conv) and  $1 \times 1$  convolution ( $1 \times 1$  Conv). The batch normalization unit BN and the nonlinear activation unit RELU are added to the output result. We compared the effects of three different feature extraction network models on the results.

### Optimize RPN

All target candidate regions can be found on the feature map through the RPN network, which is composed of convolutional layers, etc. In the RPN network of the Faster-RCNN algorithm, there are preset 9 kinds of anchors corresponding to 3 kinds of scales and 3 kinds of aspect ratios. It can improve the accuracy to choose the appropriate scale and aspect ratio for different datasets. In the infrared image dataset produced in this article, the recognition rate of some insulator strings and other objects is low due to the small area occupied in the image. To solve the problem of low recognition rate caused by the small area occupied in the image, we have added a set of scales of  $64^2$  to the RPN network and added the candidate boxes' number from 9 to 12; the results show that the recognition rate has been significantly improved.

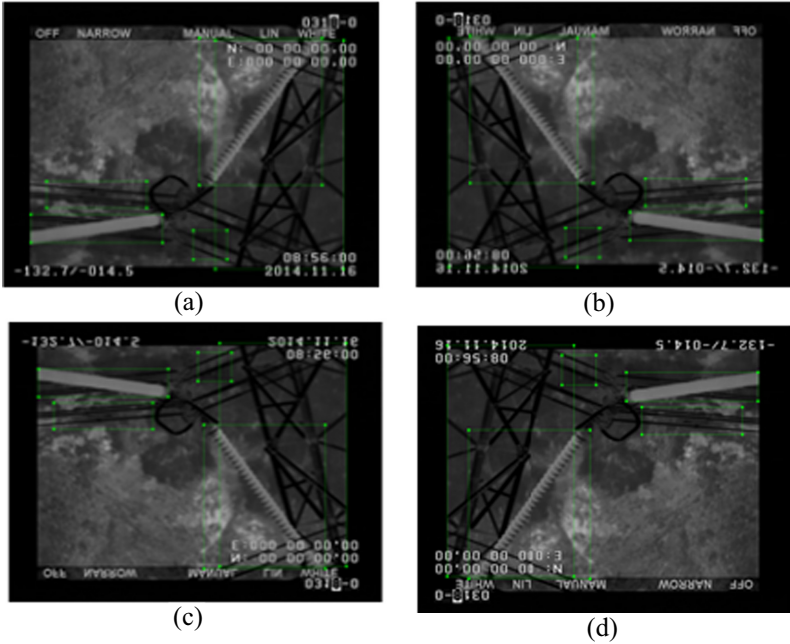
## 3 Experiment

### 3.1 Dataset Establishment

To obtain the dataset required for deep learning training, we made the dataset through the video collected by the infrared thermal imaging camera. The flow chart of our dataset's production is shown in Fig. 3. In this article, we intercepted the pictures in the transmission line video collected by the infrared thermal imaging camera, and then selected 850 clear pictures and marked the transmission lines, towers, and insulator strings in the pictures, and finally produced the VOC2007 dataset. During the labeling process, areas where the image is too blurry are not labelled. Deep learning often requires a lot of data training. To solve the insufficient of image data, we have expanded it by flipping and rotating the image. Then the data set was expanded to 3400 sheets, of which 3060 sheets were used as the training set and 340 sheets were used as the test set through data enhancement. An example of dataset annotation is shown in Fig. 4.



Fig. 3. Dataset production's flow chart.



**Fig. 4.** Dataset enhancement. (a) Original image (b) Flip left and right (c) Flip upside down (d) Rotate 180°

**Evaluation Index**

To evaluate the effectiveness of Faster-RCNN in infrared image target detection, the mean Average Precision (mAP) is used as the evaluation standard. The mAP can effectively characterize the global performance of the algorithm.

Intersection over Union (IoU) is an important index to measure the coincidence degree of different regions on the same image. For the model prediction area  $D$  and the real labeled area  $G$ , the intersection ratio represents the ratio of their intersection and union. Specifically, it can be defined as:

$$IoU = \frac{D \cap G}{D \cup G} \tag{1}$$

When the IoU of the model predicted area  $D$  and the real labeled area  $G > 0.5$ , the predicted area is considered correct. True Positive (TP), False Positive (FP), and False Negative (FN) are basic indicators commonly used in machine learning. For a certain type of target in target detection, TP represents and marks The number of prediction regions with  $IoU > 0.5$  (if there are multiple detection results that match the same labeled region, it will only be calculated once); FP means the number of prediction regions with  $IoU \leq 0.5$  in the labeled region and redundant detection results that match the same labeled region Quantity; FN represents the number of marked areas that have no matching results.

The mAP is calculated by precision (Precision, P) and recall (Recall, R). The expressions of precision rate and recall rate are as Eqs. (2) and (3).

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \quad (2)$$

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \quad (3)$$

### 3.2 Analysis of Results

#### Compare Different Convolutional Network Models

The VGG16, Resnet101, and MobileNet networks were selected for feature extraction, and different iterations were performed on the training data set. Table 1, Table 2, and Table 3 respectively represent the three network training results.

**Table 1.** Extracting network results based on VGG16 features.

Iters	Transmission line	Tower	Insulator	mAP
80000	0.899	0.9091	0.8136	0.8739
70000	0.8972	0.9091	0.8142	0.8735
60000	0.8971	0.9091	0.814	0.8734
50000	0.8854	0.9091	0.8135	0.8693
40000	0.8712	0.9091	0.8114	0.8639
30000	0.8326	0.9085	0.8057	0.8489
20000	0.7933	0.9061	0.8007	0.8333
10000	0.6229	0.9018	0.762	0.7622

**Table 2.** Extracting network results based on Resnet101 features.

Iters	Transmission line	Tower	Insulator	mAP
80000	0.8893	0.9091	0.8167	0.8717
70000	0.8891	0.9058	0.8167	0.8705
60000	0.8908	0.9071	0.8156	0.8712
50000	0.8553	0.9031	0.8134	0.8573
40000	0.8553	0.893	0.8117	0.8533
30000	0.8336	0.9091	0.7978	0.8468
20000	0.7092	0.9057	0.7978	0.8042
10000	0.6849	0.8979	0.7127	0.7652

**Table 3.** Extracting network results based on MobileNet features.

Iters	Transmission line	Tower	Insulator	mAP
80000	0.7924	0.9082	0.7772	0.8259
70000	0.7952	0.9082	0.7958	0.8331
60000	0.7853	0.9063	0.7731	0.8216
50000	0.7603	0.9037	0.7842	0.8161
40000	0.6869	0.9052	0.7155	0.7692
30000	0.6098	0.8957	0.7571	0.7542
20000	0.5845	0.8994	0.6483	0.7207
10000	0.5143	0.8704	0.6094	0.6647

The results show that the VGG16 network and Resnet101 network are better than the MobileNet network for target recognition in the transmission line. Therefore, we optimize the Faster-RCNN algorithm based on the VGG16 network and the Resnet101 network to improve the accuracy.

### Parameter Optimization

When the proportion of the recognized object in the picture is relatively low, the recognition rate of the object will be significantly reduced. To solve the problem of the decline in the recognition rate caused by the low proportion, we have added a set of  $64 \times 64$  scales to the Faster-RCNN based on VGG16 and ResNet101 to improve the recognition rate. The precision-recall curve of these networks before and after the improvement is shown in Fig. 5, and its statistical table is shown in Table 4.

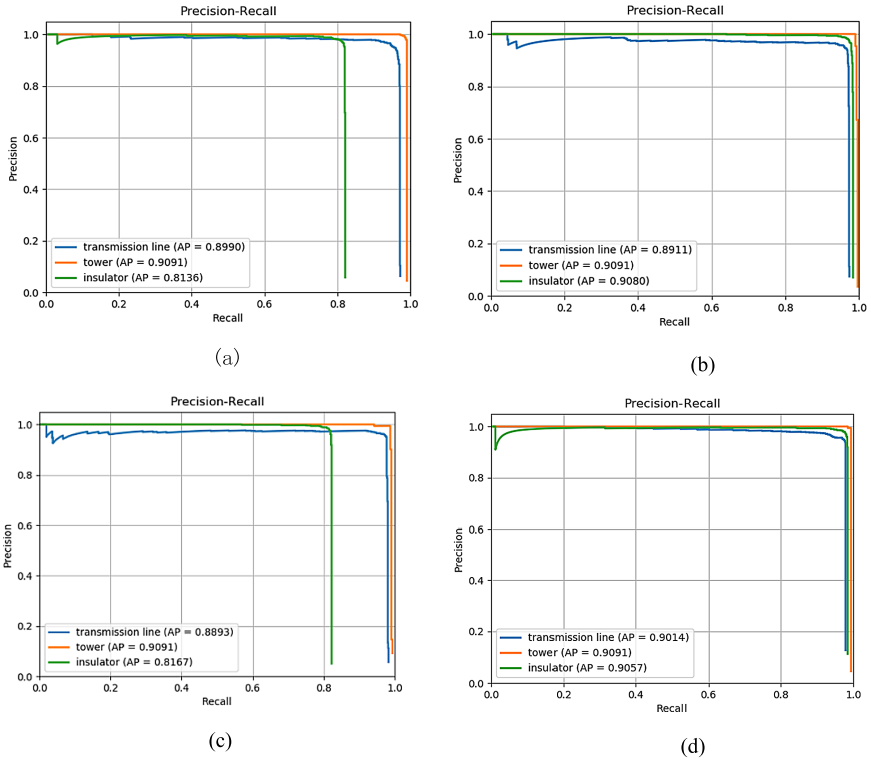
We can see that Resnet101 performs slightly better than the VGG16 model after improvement from the above table. After adding a set of scales, the accuracy of the insulator string has been greatly improved, by about 8.4%, and the overall accuracy by about 3% which showing that the improved method has certain effectiveness.

To further reflect the improvement of network fine-tuning for classification, we show the change of the loss value when training the network with the number of iterations in Fig. 6.

We set the total number of iterations to 80,000, and then conducted four sets of experiments with VGG16 and ResNET101 as the feature extraction network and controlling whether to fine-tune as a condition.

It can be seen from the figure that whether VGG16 or ResNET101 is used as feature extraction, the fine-tuned network can converge at a faster speed and converge to a lower loss value.

From the perspective of different feature extraction networks, the ResNET101 network can converge at a smaller loss value than the VGG16 network. ResNET101 is more suitable for use as a feature extraction network in this study. This conclusion is consistent with what we have obtained from Table 4.

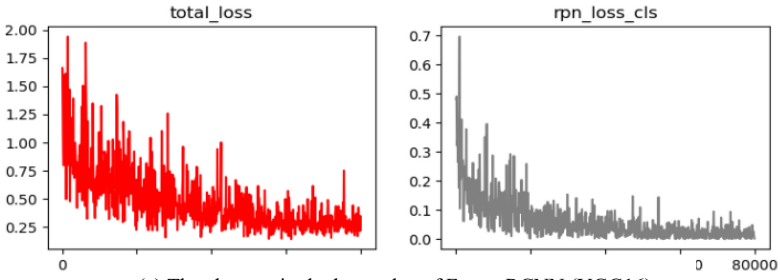


**Fig. 5.** The precision-recall graph. (a) Faster-RCNN (VGG16) precision-recall curve. (b) Improved Faster-RCNN (VGG16) precision-recall curve. (c) Faster-RCNN (ResNet101) precision-recall curve. (d) Improved Faster-RCNN (ResNet101) precision-recall curve.

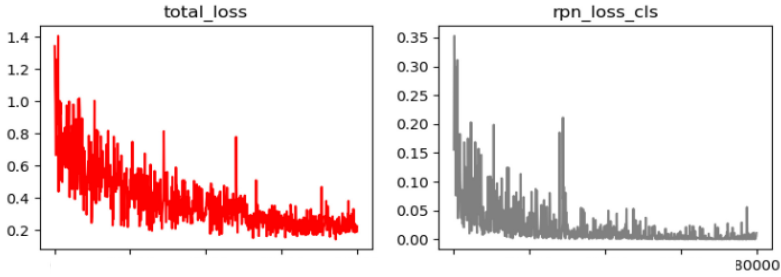
**Table 4.** Comparison before and after parameter optimization

Neural networks	Number of candidate boxes	Target detection AP value/%			mAP/%
		Power line	Tower	Insulator	
Faster-RCNN (VGG16)	9	0.899	0.9091	0.8136	0.8739
Faster-RCNN (ResNet101)	9	0.8893	0.9091	0.8167	0.8717
Faster-RCNN (VGG16)	12	0.8911	0.9091	0.908	0.9027
Faster-RCNN (ResNet101)	12	0.9014	0.9091	0.9057	0.9054

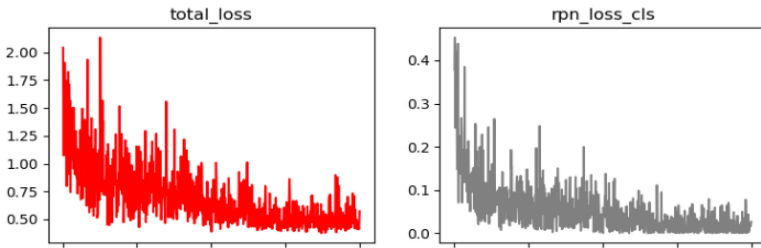




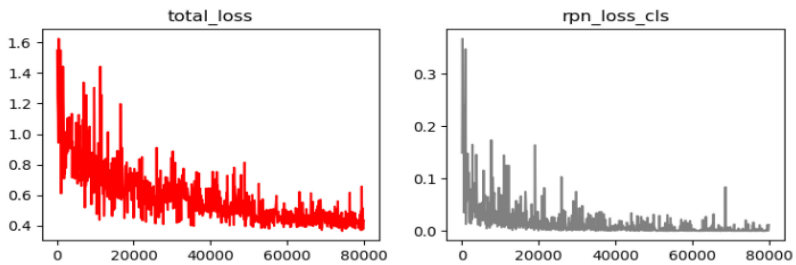
(a) The changes in the loss value of Faster-RCNN (VGG16).



(b) The changes in the loss value of improved Faster-RCNN (VGG16).



(c) The changes in the loss value of Faster-RCNN (ResNet101).



(d) The changes in the loss value of improved Faster-RCNN (ResNet101).

**Fig. 6.** The training network loss value with the number of iterations.

### 3.3 Experiment

According to the above improvement method, the Faster RCNN algorithm based on ResNet101 after adding a set of scales is finally selected and tested on the untrained data set. The testing effect is shown in Fig. 7. We can see that the algorithm accurately identifies transmission lines, insulator strings, and towers.

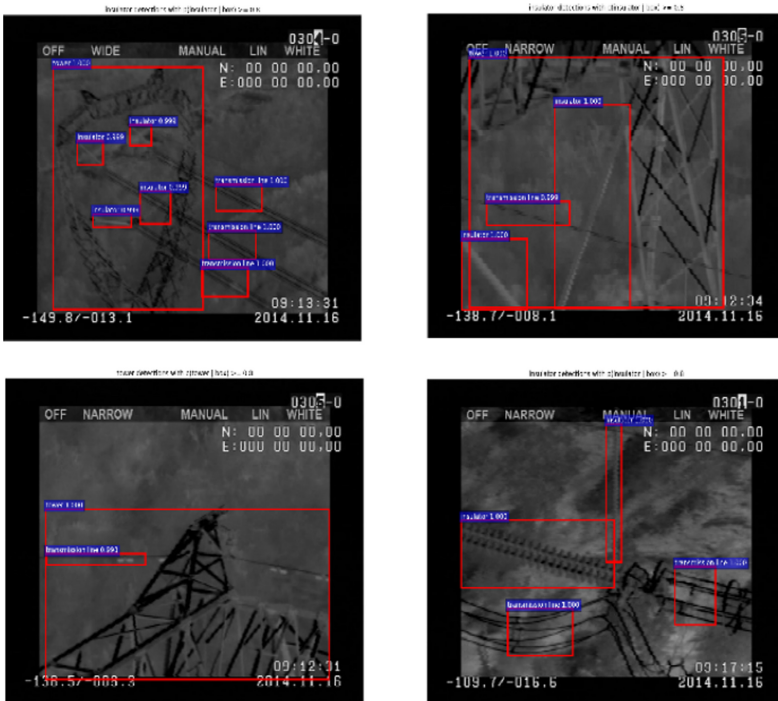


Fig. 7. The results on the test set.

## 4 Conclusion

For the difficult positioning of transmission line components under infrared image conditions, the Faster-RCNN algorithm is used to compare the target recognition effect of different feature extraction networks; we add a set of scales to solve the target area being too small.

The recognition accuracy of the insulator string is increased by about 8.4%, and the average recognition accuracy of the overall category is increased by about 3%, which verifies the effectiveness of the method. In future research, we will train the faulty equipment pictures of the transmission line and identify and locate the faulty device category and location.

**Acknowledgements.** This work was supported by the Jilin Provincial Science and Technology Department (NO.20190303016SF) and the 13th Five-year Plan for Science and Technology Project of the Education Department of Jilin Province (JJKH20170913KJ).

## References

1. Chen, W., Long, G., Yao, L., Sheng, Q.: AMRNN: attended multi-task recurrent neural networks for dynamic illness severity prediction. *World Wide Web* **23**(5), 2753–2770 (2020)
2. Chen, W., Yue, L., Li, B., Wang, C., Sheng, Q.Z.: DAMTRNN: a delta attention-based multi-task RNN for intention recognition. In: Li, J., Wang, S., Qin, S., Li, X., Wang, S. (eds.) ADMA 2019. LNCS (LNAI), vol. 11888, pp. 373–388. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-35231-8\\_27](https://doi.org/10.1007/978-3-030-35231-8_27)
3. Chen, W., et al.: EEG-based motion intention recognition via multi-task RNNs. In: Proceedings of the 2018 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics, pp. 279–287 (2018)
4. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: IEEE International Conference on Computer Vision and Pattern Recognition, pp. 580–587 (2014)
5. Girshick, R.: Fast R-CNN. In: IEEE International Conference on Computer Vision (ICCV), pp. 1440–1448 (2015)
6. Lin, G., Wang, B., Peng, H., Wang, X., Chen, S., Zhang, L.: Multi-target detection and location of transmission line inspection image based on improved Faster-RCNN. *Electr. Power Autom. Equipment* **39**(5), 213–218 (2019)
7. Liu, Y., Pei, S., Wu, J.: Deep learning based target detection method for abnormal hot spots infrared images of transmission and transformation equipment. *South. Power Syst. Technol.* **13**(2), 27–33 (2018)
8. Liu, C., Wen, Y., Hua, W., Yang, G.: Detection of power poles based on fusion of geographical location information. *J. Huazhong Univ. Sci. Technol. (Nat. Sci. Edn.)* **41**(41), 208–210 (2013)
9. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(6), 1137–1149 (2017)
10. Tang, M., Dai, L., Lin, C., Wang, F., Song, F.: Application of unmanned aerial vehicle in inspecting transmission lines. *Electr. Power* **46**(3), 35–38 (2013)
11. Tao, X., Zhang, D., Wang, Z., Liu, X., Zhang, H., Xu, D.: Detection of power line insulator defects using aerial images analyzed with convolutional neural networks. *IEEE Trans. Syst. Man Cybern. Syst.* **50**(4), 1486–1498 (2018)
12. Tong, W., Li, B., Pei, L.: Extraction and recognition of insulator based on aerial image. In: 2011 International Conference on Electric Information and Control Engineering, Wuhan, pp. 4195–4198 (2011)
13. Wang, B., Chen, L., Zhang, Z.: A novel method on the edge detection of infrared image. *Optik Int. J. Light Electr. Opt.* **180**, 610–614 (2019)
14. Wang, W., Tian, B., Liu, Y., Liu, L., Li, J.: Study on the electrical devices detection in UAV images based on region based convolutional neural networks. *J. Geo Inf. Sci.* **19**(2), 256–263 (2017)
15. Yan, S., Jin, L., Duan, S., Zhao, L., Yao, C., Zhang, W., Power line image segmentation and extra matter recognition based on improved Otsu algorithm. In: 2013 2nd International Conference on Electric Power Equipment - Switching Technology (ICEPE-ST), Matsue, pp. 1–4 (2013)
16. Yetgin, Ö., Gerek, Ö.N.: Automatic recognition of scenes with power line wires in real life aerial images using DCT-based features digit. *Sig. Process* **77**, 102–119 (2018)

17. Yue, L., Shen, H., Wang, S.: Exploring BCI control in smart environments: intention recognition via EEG representation enhancement learning. *ACM Trans. Knowl. Discov. Data (TKDD)* **15**(5), 1–20 (2021)
18. Yue, L., Tian, D., Jiang, J., Yao, L., Chen, W., Zhao, X.: Intention Recognition from Spatio-Temporal Representation of EEG Signals. In: *ADC*, pp. 1–12 (2021)
19. Yue, L., Tian, D., Chen, W., Han, X., Yin, M.: Deep learning for heterogeneous medical data analysis. *World Wide Web* **23**(5), 2715–2737 (2020)
20. Zhao, Z., Wang, L.: Aerial insulator string image automatic location method. *Chin. J. Sci. Instrum.* **35**(3), 558–565 (2014)
21. Zhu, S.: *Insulator Identification and Location of High Voltage Transmission Line Image*. Anhui University (2018)
22. Zou, D.: *Research on Recognition of Key Components and Detection of Anomaly in Transmission Line*. School of Electrical and Electronic Engineering, North China Electric Power University (2017)