Research on Recognition and Location Method of Insulator in Infrared Image Based on Deep Learning

Hua Huang1,a, Yongxi Huang1,b, Xiaojing Mu1,c, Xiaozhou Wang1,d

1China Electric Power Research Institute, Wuhan, Hubei, China

aemail: 147816514@qq.com, bemail: 648402794@qq.com
cemail: muxiaojing@epri.sgcc.com.cn, demail: 15072452923@139.com

*Corresponding author’s e-mail:muxiaojing@epri.sgcc.com.cn

Abstract. Infrared thermography technology is widely used in the thermal condition detection of insulators due to its advantages of non-contact, sensitive, online detection. To realize the automatic detection of the operating condition of insulators in complex environments, this paper proposes a method for the recognition and location of the insulator based on Region-based Fully Convolutional Networks (R-FCN). The model was trained and tested on the constructed insulator infrared data set, compared with the SSD model. The results showed that the R-FCN detecting insulators can not only accurately locate insulators, but have an AP (average precision) value as high as 89.2%. Therefore, the findings in this paper have verified that R-FCN has great advantages in the recognition and location of infrared images of insulators and has practical application value.

1. Introduction

As a widely distributed and numerous fitting, the insulator can achieve mechanical fixation and electrical insulation functions, thereby playing an important role in maintaining the safe and stable operation of the electric power line. Once the insulator fails, the power supply reliability of the power grid will be reduced, which may cause a large-area blackout and great economic loss in serious cases [1]. To ensure the safe and reliable operation of the power grid, it is important to detect the operating condition of insulators in time.

Infrared thermography technology can show the infrared heat radiated by the target device in the form of temperature through non-contact shooting. It has the advantages of high sensitivity and immunity to electric field interference and is widely used in the power industry. However, the detection of power infrared image data mostly relies on the recognition and analysis of experienced power engineers [2]. This method will consume a lot of manpower and time costs, severely reduce the efficiency of power equipment condition detection, and cannot meet the needs of intelligent operation and inspection.

As a possible solution, the object detection method based on deep learning can improve the accuracy and speed of insulator recognition and location, and then realize the intelligent analysis and processing of insulator image information. Thus, it has become a study hotspot in recent years. In [3], BRISK was used to construct insulator features, and the deep convolution feature map aggregation method was used to recognize infrared insulators. To improve the location accuracy of insulators, the SVM classifier is embedded in a multi-scale sliding window frame. Reference [4] applied the SSD [5] deep learning model and Field Programmable Gate Array hardware acceleration method to the power inspection UAV and
designed an insulator front-end location system based on image recognition technology to realize the insulator real-time automatic location. Reference [6] used Faster R-CNN [7] to recognize and locate insulators and studied the recognition of insulator self-explosion defects on this basis. The results showed that this method can maintain high accuracy and robustness even under complex background. The above methods have laid an advanced theoretical foundation for solving the problem of insulator automatic detection in the power grid and provided a new idea for studying the recognition and location of infrared insulator images with complex backgrounds and low resolution. Based on this, this paper proposes a recognition and location method for insulators based on R-FCN [8] to achieve accurate detection of infrared insulator images in complex environments.

2. R-FCN network architecture and principles
In this paper, R-FCN deep learning algorithm is used to realize the recognition and positioning of infrared images of insulators. This section will discuss the R-FCN network architecture and principles in detail.

2.1. R-FCN network architecture
The R-FCN model is composed of a fully convolutional network, and its network architecture mainly includes a feature extraction network (ResNet), a region proposal network (RPN), and a position-sensitive region of interest (ROI) detection network. Detailed architecture as shown in Figure 1.

In order to ensure that the network structure is still translation invariant after the introduction of a position-sensitive ROI detection network, R-FCN adds a position-sensitive score map (PSM) layer after the feature extraction network and the RPN network [9]. Firstly, R-FCN inputs the deep features of image extracted by ResNet and the object candidate region generated by RPN into the PSM layer. And then it uses the position-sensitive ROI pooling layer in the ROI detection network to classify the candidate objects and regress the frame.

2.2. ResNet feature extraction network
ResNet, was proposed by Microsoft Research, won the championship in the ILSVRC2015 competition with a top-5 error rate of 3.57%. Different from the previous feature extraction network, ResNet introduces the idea of cross-layer connection and builds a residual module. It not only greatly expands the depth and width of the network, but also accelerates the training of deep neural networks very quickly. On this basis, the accuracy of the model is further improved. The core component residual module is
shown in Figure 2. It can be seen that the module has two branches, namely the short-circuit connection branch and the convolution transfer branch. The convolution branch consists of two convolution layers and an activation function layer. \( x \) represents the input data of the residual module, \( W \) represents the weight parameter of each convolution module in the residual unit, and \( f(x) \) represents the residual map that needs to be learned.

\[
\begin{align*}
&x_{\text{input}} \\
&\text{First weight layer } W_1 \\
&\text{ReLU} \\
&\text{Second weight layer } W_2 \\
&+ \\
&\text{ReLU} \\
&f(x_{\text{input}}) \\
&f(x_{\text{input}}) + x_{\text{input}}
\end{align*}
\]

Figure 2. Residual module structure

2.3. Region proposal network
RPN based on Convolutional Neural Network was first proposed in Faster-RCNN. It takes the convolution feature map as input and is used to generate candidate regions of the object. The principle is shown in Figure 3.

After the feature map is input into the RPN, a set of candidate boxes and corresponding scores will be generated. A sliding window of \( n \times n \) size is selected for the feature map, and each input window will predict the candidate regions where \( k \) insulators may appear, which are called anchors. Each anchor is composed of anchor points with aspect ratio, size, and sliding window properties. Among them, the sliding window takes the anchor point as the center and its size is \( 3 \times 3 \), which corresponds to the frames of three proportions and three sizes in the original image. In addition, multiple similar target candidate boxes may appear for the same target in the positioning process. In order to make a object finally retain
only one candidate box, it is necessary to use non-maximum suppression (NMS) to delete redundant boxes. NMS adjusts the anchor frame according to the given intersection ratio score. After calculating the convolution of the anchor box, the offset information of $4k$ prediction candidate boxes will be obtained in the Region Layer. Among them, it includes the position $x$ and $y$ of the candidate frame center point and the width $h$ and height $w$ of the candidate frame; $2k$ foreground and background classification score information is obtained in the Class Layer. Therefore, with the help of the spatial relevance of the RPN to the convolution feature map, the negative sample area can be quickly eliminated, and the object candidate area can be generated synchronously.

### 2.4. ROI detection network

R-FCN adds a position-sensitive score map (PSM) convolutional layer to the last layer of the feature extraction network, which outputs a PSM with a dimension of $k^2(C+1)$ ($C$ represents the number of target categories, and 1 represents the background). In order to improve the calculation speed of the model, the fully connected layer in the ROI detection network is replaced by a position-sensitive ROI pooling layer, which performs a pooling operation on the PSM. In the final detection process, $k^2$ score maps are generated for each category. R-FCN divides each ROI rectangular box into $k \times k$ grid areas. For a rectangular box with a size of $w \times h$, the size of a grid is approximately equal to $(w/k)(h/k)$. The $C$ category pooling operation for the $(i, j)$th grid of the ROI is as follows:

$$r_c(i, j|\theta) = \sum_{(x, y) \in \text{label}(i, j)} z_{i, j, c}(x + x_0, y + y_0|\theta) / n$$  \hspace{1cm} (1)$$

Among them, $(x_0, y_0)$ is the coordinates of the upper left corner of the ROI, $\text{bin}(i, j)$ represents the interval range of the grid, $z_{i, j, c}$ represent the newly generated score map, $\theta$ is the network learning parameter, and $n$ is the number of pixels in each grid.

After the pooling process is completed, vote on the ROI. Sum the $k \times k$ grids to get the $(C+1)$ dimensional output, which is the score of the category judgment. Then classify through the softmax layer, and the output result at this time is the position coordinates and category name of the insulator.

### 3. Experiment and result analysis

In order to evaluate the effectiveness and advantages of the model, the R-FCN model was trained and tested on the constructed insulator infrared data set, and the results were compared with the SSD model.

#### 3.1. Experimental description

The original infrared image used in this study comes from a power grid company. The experiment expands the data set through operations such as rotation, symmetry, and color gamut transformation, thereby enhancing the diversity of the data. The constructed insulator infrared data set has a total of 2760 images, 60% of the data set is used as the training set, and the remaining 40% is used as the test set. The entire experiment runs on a server with 1080Ti GPU, and uses the TensorFlow deep learning framework to train and test the R-FCN model.

In this paper, the average precision (AP) of the object detection accuracy evaluation index is used to evaluate the model, such as formula (2). AP value is a measure of the object detection model's ability to predict the position and category of a single object. The detection threshold, intersection over union (IoU), needs to be set before calculating the AP. IoU refers to the area ratio of the intersection of the prediction box and the label box to the union.

$$AP = \int_0^1 \rho(r) dr$$  \hspace{1cm} (2)$$

#### 3.2. Training process and analysis

At the beginning of training, transfer learning is used to initialize the weight of the model, so that the model can still achieve good training results in small sample learning tasks. In addition, the batch size of the model is set to 8, and the number of iterations is 120000. The loss curve reflects the difference between the model's predicted value and the true value. The smaller the loss value, the stronger the
model’s anti-interference ability. When the training loss converges to near the lowest value, the model is stopped training and saved. The loss curve of the R-FCN model trained on the insulator infrared data set in this paper is shown in Figure 4. It can be seen that as the number of training iterations increases, the loss curve gradually decreases with certain fluctuations. When the training iteration reaches about 100000 steps, the R-FCN loss value has stabilized, and the robustness is better.

![R-FCN loss curve](image)

**Figure 4. R-FCN loss curve**

### 3.3. Test results and comparative analysis

The trained R-FCN model was tested on the constructed insulator infrared test set and compared with the mainstream object detection model SSD. The test results are shown in Table 1. It can be seen that when the IoU threshold is 0.5, the detected AP value of the R-FCN model is as high as 89.2%, which is 4.2% higher than that of the SSD model. It proves the advantage of R-FCN model in detection performance.

| Methods | AP@0.5 (%) |
|---------|------------|
| R-FCN   | 89.2       |
| SSD     | 85.0       |

Figure 5 shows the visualization results of the R-FCN model on some images of the test set, which includes the category name, prediction frame, and confidence of the insulator. It can be seen that the model can not only identify the insulator with high accuracy, but also accurately locate the insulator position. It laid the foundation for the condition evaluation and intelligent diagnosis of the insulator in the next step.

![Visual detection results of R-FCN on some images of the test set](image)

**Figure 5. Visual detection results of R-FCN on some images of the test set**
4. Conclusion

In this paper, R-FCN model based on the recognition and location of insulators in infrared images is constructed. And the principle of the R-FCN model is introduced from the perspective of network structure composition. Transfer learning is used to improve the model learning efficiency. R-FCN model is experimentally verified under the TensorFlow deep learning framework and compared with the SSD model. The results show that the average accuracy of R-FCN is as high as 89.2% on the constructed insulator infrared data set, 4.2% higher than that of SSD. And also, it can accurately locate the insulator position. Therefore, the findings illustrate that the constructed R-FCN model can be used to accurately recognize and locate the insulator in the infrared image, which lays a solid foundation for further evaluation and diagnosis of the insulator condition.

References

[1] Huang, C., Li, Y. (2014) Analysis on flashover characteristics based on pollution flashover model of composite insulator. Insulators and Surge Arresters, 4: 31-35.
[2] Zheng, H., Li, J., Liu, Y., et al. (2021) Infrared object detection model for power equipment based on improved YOLOv3. Transactions of China Electrotechnical Society, 36: 1389-1398.
[3] Zhao, Z., Xu, G., Qi, Y. (2016) Representation of binary feature pooling for detection of insulator strings in infrared images. IEEE Transactions on Dielectrics and Electrical Insulation, 23: 2858-2866.
[4] Su, C., Yang, Q. (2019) Insulator Front-end Location Based on Computer Vision Deep Learning. Guangdong Electric Power, 32: 33-40.
[5] Liu, W., Anguelov, D., Erhan, D., et al. (2016) SSD: Single shot multibox detector. European conference on computer vision. In: European Conference on Computer Vision. Cham. 21-37.
[6] Guo, T., Yang, H., Shi. L., et al. (2019) Self-Explosion Defect Identification of Insulator Based on Faster Rcnn. Insulators and Surge Arresters, 3:183-189.
[7] Ren, S., He, K., Girshick, R., et al. (2016) Faster R-CNN: towards real-time object detection with region proposal networks. IEEE transactions on pattern analysis and machine intelligence, 39: 1137-1149.
[8] Dai, J., Li, Y., He, K., et al. (2016) R-FCN: object detection via region-based fully convolutional networks. In: Neural Information Processing Systems. Barcelona. 379–387.
[9] Fan, X., Shen, C., Xu, J., et al. (2019) A method of Traffic Sign Recognition Based on Cascade and End-to-end Depth Architecture. Computer Measurement & Control, 27: 143-148.