Research on Construction of Infrared Image Classification Model of Substation Equipment Based on CNN

Kehui Zhou, Zhiwei Liao and Xiaochun Zang

Abstract Infrared Intelligent detection is the key research direction of infrared fault detection of power equipment. However, different types of equipment and different processing purposes lead to a variety of fault detection methods. If the manual method is used to classify the equipment, the efficiency of fault detection will be greatly reduced. In this paper, based on convolutional neural networks, based on RGB and HSV color space conversion, a classification model suitable for infrared images of power equipment is constructed. Firstly, the structure characteristics and training process of CNN are introduced. After that, based on RGB and HSV color space conversion, the infrared image of substation equipment is processed, and the target area of suitable size is extracted to establish network training and test set. Finally, a CNN-based infrared image classification model is established, and its good applicability is verified by case analysis.

Keywords Substation equipment · Infrared detection · CNN · Image processing · Image classification

1 Introduction

As an essential element of power system operation, electrical equipment analysis is a powerful guarantee for the stable operation of power systems [1]. Infrared detection technology based on the principle of infrared radiation has become the necessary means for daily maintenance of electrical equipment with the advantages of non-contact, high efficiency and intuitive image. It plays an important role in the state detection of devices such as arresters, circuit breakers and insulator strings [2–4]. However, the current technology also has the disadvantages of excessive human factors and low intelligence. With the rise of inspection robots and intelligent substations [5], infrared detection technology based on image processing and

K. Zhou (✉) · Z. Liao · X. Zang
School of Electric Power, South China University of Technology, Guangzhou, China
e-mail: 2903807982@qq.com

Y. Xue et al. (eds.), *Proceedings of PURPLE MOUNTAIN FORUM 2019-International Forum on Smart Grid Protection and Control*, Lecture Notes in Electrical Engineering 585, https://doi.org/10.1007/978-981-13-9783-7_84

machine learning has become a development trend. The literature [6, 7] propose to use the partial feature aggregation description vector as the fully connected layer of the deep learning network to mark the insulator string in the infrared image. In [8], the temperature histogram of the infrared image of the arrester is proposed to extract the characteristic quantities such as the highest temperature, average temperature and temperature standard deviation, which are used as part of the neural network input to determine the state of the leakage current of the arrester.

From the above literature analysis and summary, different types of equipment and different processing purposes lead to diversity in method selection, both in image processing method selection and in machine learning method selection. If selecting the processing method flow according to the traditional manual classification result is still adopted, the infrared detection process will be further complicated. Therefore, it is imperative to introduce image classification into the infrared detection, which is helpful to greatly improve the intelligence degree. The image classification method based on Convolutional Neural Network (CNN) has gained extensive attention, research and application due to its excellent performance in multiple image classification competitions [9, 10]. In [11], based on the characteristics of each layer of convolution-al neural network, an insulator detection method based on cross-connected convolutional neural network is proposed. In [12] proposed a method for image recognition of power equipment combined with deep learning and random forests to solve the problem of intelligent analysis and identification of massive unstructured image data. Based on the convolutional neural network, this paper explores and constructs a classification model suitable for infrared images of power equipment. Firstly, it introduces the structural characteristics and training process of CNN. Then, an infrared image preprocessing method for equipment classification is proposed. Based on this, a network of the training and test is established. Finally, the good applicability of the CNN-based infrared image classification model is verified from the classification accuracy and classification efficiency by examples.

2 The Structure and Principle of CNN

At the end of the last century, foreign scholars proposed a character recognition model based on CNN and achieved good results in handwritten digit recognition [13]. Then, with proposal of deep learning concept, the expansion of database scale and hardware development, CNN models with better recognition and deeper levels emerge in an endless stream. Overall, the basic structures of the network are still composed by input layers, convolution layers, excitation layer, pooling layer, fully connected layer and the output layer.

The CNN realizes image feature extraction through the alternating of the convolution layer and the pooling layer, wherein the convolution layer convolves the input image or the output of the upper layer through a learnable convolution kernel to obtain a feature map. Each convolution kernel can convolve a combination of multiple feature maps, which is calculated as follows:

$$c_j^l = \sum_{i \in M_j} x_i^{l-1} \otimes k_{ij}^l + b_j^l \tag{1}$$

Wherein, c_j^l is the value of the l convolutional layer j channel, which is obtained by convolving the $l-1$ layer feature map; M_j is a c_j^l calculated feature atlas, which can be a combination of l layer feature maps; k_{ij}^l is a convolution kernel whose number determines the output characteristics of the layer; b_j^l is the offset. For each l layer output feature map, the convolution kernel corresponding to the calculated l layer feature map may be different, and $\otimes$ represents the convolution calculation. The formula is as follows:

$$S_{ij} \otimes K_{ij} = \sum_{i=1}^{m} \sum_{j=1}^{m} S_{ij} K_{ij} \tag{2}$$

Wherein, m is the convolution kernel size; K is the convolution kernel value, S is the input image or the feature map value. The convolution kernel traverses the entire image in a certain step size s to obtain a convolution result for the image.

Each feature map in the CNN greatly reduces the network parameters through the weight sharing and partial perception of the convolution kernel. It extracts the advanced features of the image through the combination of features from the bottom layer to the upper layer. The output of the convolutional layer is nonlinearly mapped by the excitation layer, and ReLU is the commonly excitation function in CNN. The pooling layer, also known as the lower layer, is usually located in the middle of a continuous convolution layer, which can reduce features and ensure partial invariance of features [14]. Common pooling methods are maximum pooling and average pooling. The maximum pooling usage is higher in reality. As shown in Fig. 1, the same color area is a pooled area (taking 2 × 2 cores as an example), and the maximum value in the area is taken as the output result and reconstitute the feature map.

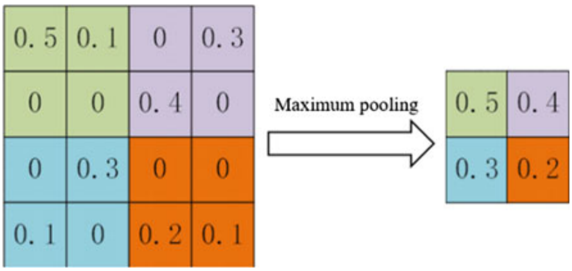

Fig. 1 The maximum pooling method

The fully connected layer is usually located in the upper layer of the CNN and is connected to each neuron in the previous layer. The weighted summation of the input feature map is spliced into a one-dimensional feature, which is equivalent to convolution with the same convolution kernel as the size and size of the previous layer, which is used to extract the advanced features of the image as input to the classifier. For classification tasks, the output layer is usually a classifier, with softmax being the most common [15].

CNN training is similar to traditional neural networks, used Gradient descent algorithm and backpropagation algorithm. The training process consists of forward propagation of signals and reverse propagation of errors. For forward propagation, the weights W and offsets b are first randomly initialized, and then the output values H of each layer are calculated based on the initialization data:

$$H = F(WX + b) \tag{3}$$

Wherein, H is the calculated output value for each layer; W is the weight matrix; b is the offset; F is the excitation function. The input data is calculated and transformed by each layer to obtain the final output layer data Y.

In the case of backpropagation, the difference E between the actual output and the expected output is calculated by constructing the error function. The form of E is related to the selected error function type and the gradient descent method. The commonly used error function is the sum of squared error functions. In the case of using small batch gradients:

$$E = \frac{1}{k}\sum_{i=1}^{k}\frac{1}{2}\sum_{j=1}^{m}(y_{ij} - t_{ij})^2 \tag{4}$$

Wherein, k is the number of samples in each mini-batch; m is the number of output neurons; t is the expected output. Among them, the small batch gradient descent method refers to dividing all samples equally into multiple Mini-batch, and each parameter updating process uses one of them to calculate, which has the advantages of both speed and accuracy. Update the weight and offset by gradient descent so that the output of the network reaches or approaches the desired output. The update formula is:

$$\begin{aligned} W^{(i+1)} &= W^{(i)} - \eta\frac{\partial E}{\partial W} \\ b^{(i+1)} &= b^{(i)} - \eta\frac{\partial E}{\partial b} \end{aligned} \tag{5}$$

Wherein, η is the learning rate [16].

3 Infrared Image Background Separation and Component Area Extraction

The number of devices in the substation is large and the distribution is dense. As a result of infrared image acquisition, the window not only contains the target equipment at the center position, but also includes many other surrounding equipment and supporting steel frame and other background interference, as shown in Fig. 2. The infrared image of the target device contains interferences such as current transformers, isolation switches and steel frames, which seriously affect the state judgment of the target equipment. If the original infrared image is directly applied for classification and judgment, it not only has high requirements on hardware such as computing resources, but also the accuracy of classification cannot be guaranteed. Therefore, with reference to the current infrared detection process, this paper first preprocesses the infrared image to improve the accuracy of classification and lay the foundation for the intelligent judgment of the subsequent device state.

3.1 Background Separation Based on RGB and HSV Color Space Conversion

In order to better extract the target area from the original image, the conversion of the RGB color space and the HSV color space is realized. The RGB color space is based on three basic colors of red, green and blue, and forms a rich color expression

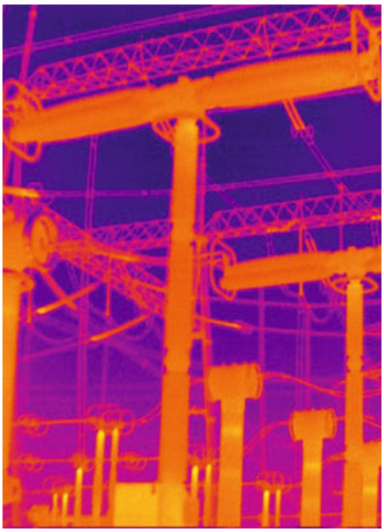

Fig. 2 Original infrared image

by superimposing different components; The HSV color space is based on hue, saturation and value. The saturation refers to the degree of close to the spectral color, and the value ranges from 0 to 1. The larger the value, the higher the saturation, the darker the color, and the closer to the spectral color. The conversion relationship between the RGB color space and the saturation S in the HSV color space is:

$$S = \begin{cases} 0 & \max = 0 \\ 1 - \min/\max & \max \neq 0 \end{cases} \tag{6}$$

Wherein, $\max = \max(R, G, B)$, $\min = \min(R, G, B)$.

In the RGB color space, the R component is [255, 0, 0]. According to (6), it converts to the red saturation dimension value S = 1 in the HSV space. According to this feature, a number of vertices are selected in the original image, respectively marked and connected in red. The boundary is extracted and the enclosed area is filled and the unrelated interference is eliminated to obtain the target area, as shown in Fig. 3.

Fig. 3 Target area extraction

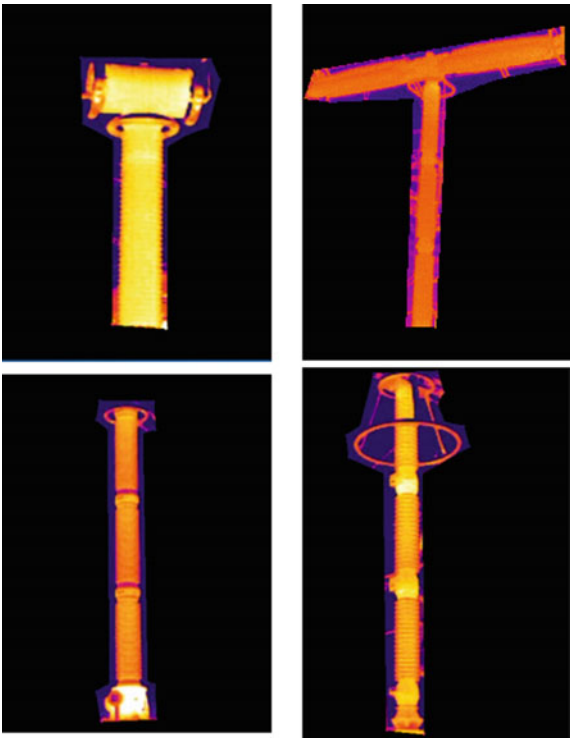

Fig. 4 Target area scaling

3.2 A Subsection Sample Target Area Scaling

Generally speaking, due to the different models of the infrared camera, the size of the infrared image captured is not the same. To unify the input image size and save computing resources, it is necessary to scale the target and retain the target information to the greatest extent. In this paper, we first detect the start and end rows, columns ($r_{\min}$, $r_{\max}$, $c_{\min}$ and $c_{\max}$) of the target area in Fig. 3, and extract the minimum rectangular outline of the target device, and then extend the image dimension to the adjacent square area with the center of the rectangular outline as the origin. Taking the image size of the MNIST dataset and the CIFAR10 dataset as a reference, this paper scales the extracted image of the target area of 300*300 to the size of 32*32 dimension, as shown in Fig. 4.

4 The Structure and Principle of CNN-Based Infrared Graphics Classification

4.1 Construction of a Training Set

Based on the image preprocessing method of Sect. 2, the original infrared images acquired in the substation are extracted and adjusted to appropriate sizes, which lay foundation to expand the scale of the training set, improve the applicability of the model, perform enhanced processing such as rotation and translation on the processed image and form a training set. This paper considers the classification of 7 types of 500 kV key equipment in substation, including arrester, circuit breaker, current transformer, voltage transformer, insulator string, isolation switch and high voltage casing. There are 1500 sheets of each training picture, a total of 10,500 pictures for training. Parts of the sample image is shown in Fig. 5.

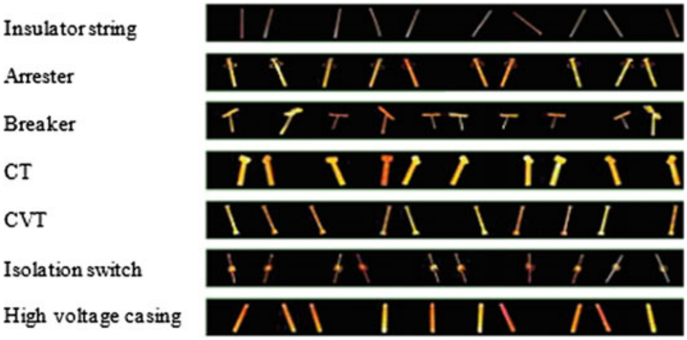

Fig. 5 Parts of the sample image

4.2 Network Construction

The infrared image classification CNN structure designed is as follows:

(1) Input layer-Input: input 32*32*3 the infrared image of the devices;
(2) Convolution layer-Conv1: the layer has 32 convolution kernels of 5*5, expanding the boundary of the input image in each direction by 2 pixels (Padding = 2) and moving step length is 1;
(3) The pooling layer-Pooing1: the layer adopts the maximum pooling; the pooling area size is 3 × 3 and the moving step size is 2;
(4) Convolution layer-Conv2: the layer has 32 convolution kernels of 5*5; Padding = 2 and the convolution kernel moves by 1 step;
(5) The pooling layer-Pooing2: the layer adopts the maximum pooling, the pooling area size is 3 × 3, and the moving step size is 2;
(6) Convolution layer-Conv3: the layer has 64 convolution kernels of 5*5; Padding = 2 and the moving step is 1;
(7) The pooling layer-Pooing3: the layer adopts the maximum pooling, the pooling area size is 3 × 3, and the moving step size is 2;
(8) Fully connected layer-FC1: output 64 one-dimensional features;
(9) Fully connected layer-FC2: output 7 one-dimensional features, which is the number of classifications;
(10) Output layer-Output.

There is a ReLU excitation layer between each convolution layer and the pooling layer. The network structure is shown in Fig. 6.

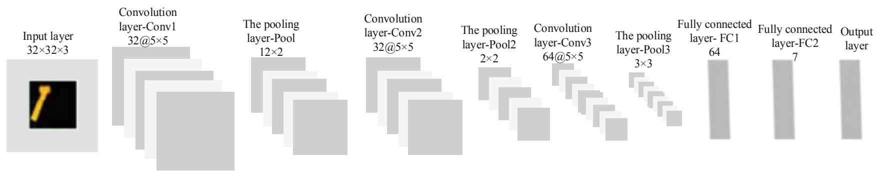

Fig. 6 The network construction

4.3 Network Training

This paper completes the construction, training and testing of CNN models on the MATLAB platform. MATLAB fully updated the neural network toolbox in 2017b version, which is more convenient for deep learning research and application from model construction, parameter setting and hardware support. As described above, after completing the CNN network structure construction, this paper uses the small batch gradient descent algorithm to train on the GPU, where the Min-batch size is set to 128, each Epoch contains 82 iterations, and the maximum iteration Epoch is set to 40. The training platform parameters are shown in Table 1.

After repeated adjustments to select appropriate training parameters, the iterative process of each step of the training process is shown in Fig. 7. As can be seen from the Fig. 7, the network training process is faster. After 20 steps of iteration, the recognition rate of each Mini-batch reaches 100%.

4.4 Case Analysis

After completing the network training, the selected part of the infrared image taken from the 500 kV substation is preprocessed to form a test set. Testing the network training result and the result is shown in Table 2.

From Table 2, the infrared image classification CNN of the substation key equipment designed and trained in this paper has a high recognition accuracy rate.

Table 1 The training platform parameters

Item	Parameter
System version	Windows 7 Ultimate SP1
CPU model	Intel(R) Core (TM) i5-4590
Memory	8G
GPU model	NVIDIA GTX750
Display memory	1G

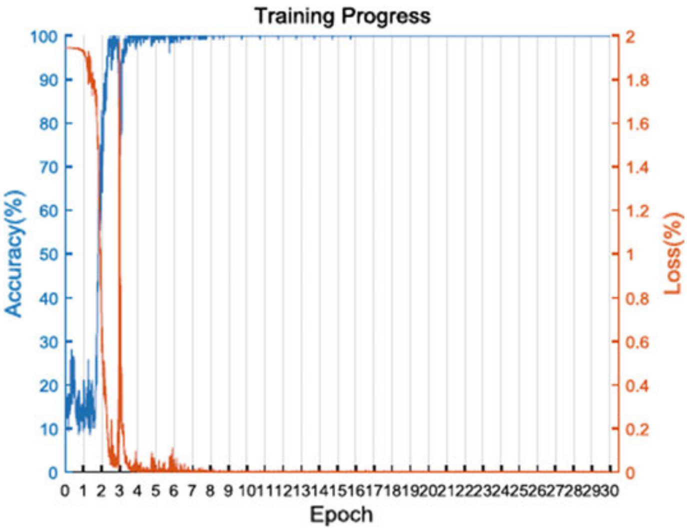

Fig. 7 Training progress

Table 2 The network training result

Equipment type	Insulator string	Arrester	Circuit break	CT
Number	693	810	732	810
Error recognition	0	1	2	0
Total time (s)	6.99	10.67	11.16	12.64
Average time (s)	0.010	0.013	0.015	0.015
Recognition rate (%)	100	99.9	99.7	100
Equipment type	CVT	Isolation switch	High voltage casing	
Number	735	243	96	
Error recognition	0	0	0	
Total time (s)	10.54	3.44	1.30	
Average time (s)	0.014	0.014	0.014	
Recognition rate (%)	100	100	100	

On the test set of 4000 images, there are only three misidentification samples, which has achieved a recognition rate of 99.9%. It has a high operational efficiency comparable to that of the manual and is fully applicable to the level of practical work. It lays a solid foundation for the subsequent intelligent diagnosis of key equipment infrared images.

5 Conclusion

In this paper, the convolutional neural network is applied to the substation infrared detection work to assist and promote the development of intelligent state judgment based on infrared equipment. Aiming at the infrared image of the complex background in the substation, an image preprocessing method for target device extraction is proposed. And based on this method, the image database of 7 type of infrared devices including insulator string, arrester and circuit breaker are constructed. What's more, CNN structure for image classification is proposed. The test results show that the network structure has high recognition accuracy and operational efficiency.

With the gradual deepening of the standardization of infrared detection work, the way of high-quality infrared image acquisition is more diversified. The introduction of intelligent methods in the process of mass data processing is a research hotspot and development trend. With strong learning and generalization capabilities, CNN is widely favored in the field of computer vision. With the increasing and perfect computing resources and large-scale databases, its application in infrared detection and even power systems will be more extensive and in-depth.

References

1. Li G, Zhang B, Zhao WQ et al (2018) Data Science issues in state evaluation of power equipment: challenges and prospects. Autom Electr Power Syst 42(21):10–20
2. Jin LJ, Tian ZR, Gao K et al (2016) Discrimination of insulator contamination grades using information fusion of infrared and visible images. Proc CSEE 36(13):3682–3691
3. Lu ZM, Liang JC, Wang TZ et al (2015) Application of SF6 infrared imaging leakage detection technology in live detection for UHV. Insulating Mater 48(09):29–33
4. Zhang JG (2015) Application of infrared temperature measurement technology to live detection of zinc oxide arrester. High Voltage Apparatus 51(6):200–204
5. Chen AW, Yue QM, Zhang ZY et al (2012) An image recognition method of substation breakers state based on robot. Autom Electr Power Syst 36(06):101–105
6. Zhao Z, Xu G, Qi Y (2016) Representation of binary feature pooling for detection of insulator strings in infrared images. IEEE Trans Dielectr Electr Insul 23(5):2858–2866
7. Zhao ZB, Fan XQ, Xu GZ et al (2017) Aggregating deep convolutional feature maps for insulator detection in infrared images. IEEE Access 5:21831–21839
8. Abdul-Malek Z (2017) Electrical and temperature correlation to monitor fault condition of Zno surge arrester. In: International conference on information technology, computer, and electrical engineering, IEEE, pp 182–186
9. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: International conference on neural information processing systems. Curran Associations Inc., New York, pp 1097–1105
10. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale imaged recognition [2014-4-20]. https://arxiv.org/abs/1409.1556
11. Zuo GY, Ma L, Xu CF et al (2019) Insulator detection method based on cross-connected convolutional neural network. Autom Electr Power Syst 43(04):101–108

12. Li JF, Wang QR, Li M (2017) Electric equipment image recognition based on deep learning and random forest. High Voltage Eng 43(11):3705–3711
13. LeCun Y, Bottou L, Bengio Y et al (1998) Gradient-based learning applied to document recognitio. Proc IEEE 86(11):2278–2324
14. Chang L, Deng XM, Zhou MQ et al (2016) Convolutional neural networks in image understanding. ACTA Automatica Sinica 42(9):1300–1312
15. Wei D, Gong QW, Lai WQ et al (2016) Research on internal and external fault diagnosis and fault-selection of transmission line based on convolutional neural network. Proc CSEE 36:21–28
16. Zhang GX, Hu Z (2011) Improved BP neural network model and its stability analysis. J Central South Univ (Sci Technol) 42(01):115–124