Insulator Self-explosion Fault Detection Based on Transfer Learning

Huohua Li[1, *], Xiaogang Fu[2], Tairu Huang[1]

[1]College of Electrical Engineering, Shanghai DianJi University; Shanghai, China

[2]Industrial Technology Center, Shanghai DianJi University, Shanghai, China

[*]Corresponding author: 3207770397@qq.com

Keywords: Fault detection, AlexNet, transfor learning, pre-training network, vgg16.

Abstract: Insulator is an important device used to support wire and electrical isolation in power system. It is of great significance to realize intelligent detection of insulator state for stable and reliable operation of transmission line. Considering the randomness of insulator image state and the complexity of image background, as well as the small scale of insulator fault dataset collected at present, the classic AlexNet network structure is built, and the trained model is in the state of under fitting, which cannot meet the requirements of detection accuracy. In order to solve this problem, the transfer learning method is introduced and the fine tuning strategy of the pre-training network is elaborated. In the experiment, based on the pre training network of vgg16, some weight parameters of the vgg16 network are frozen, and the self-defined full connected layer is trained. The convergence speed and detection accuracy of the new model in the test set and verification set are effectively improved.

1. Introduction

Since 2012, the AlexNet convolutional neural network model with eight layers network structure proposed by krizhevsky et al, which has performed excellently on the ImageNet dataset [1]. Convolutional neural network (CNN) has become the research hotspot of various research institutions. Various more advanced network structures have been proposed successively. The deep learning method based on convolutional neural network has gradually penetrated into all aspects of production and life.

Since convolutional neural network can automatically extract abstract features, computer vision technology has developed from traditional image recognition to a new stage, and has been widely used in many fields, and plays an important role in power system automation. As an important part of high-voltage overhead transmission line, insulator plays a dual role in supporting conductor and preventing current grounding. However, due to the long-term exposure of overhead lines to harsh outdoor conditions, and most of them are distributed in remote mountainous areas [2]. Affected by climate, pollution and various natural disasters, insulators will suffer from aging, self-explosion, string dropping and other faults [3], resulting in poor insulation performance or even complete loss of insulation. According to statistics, the trip accidents caused by insulator failure account for 81.3% of the current transmission line faults [4]. If the health status of insulators is detected irregularly, it will pose a hidden danger to the reliable operation of transmission lines.

At present, some fault recognition algorithms based on convolutional neural network have been applied to insulator fault detection and achieved good results. However, due to the small scale of the available data sets, the training model is difficult to improve the detection accuracy. Using data enhancement method to improve the scale of existing insulator data sets can achieve certain results, but it still can not overcome the defect of long training time of large data sets, and can not solve the model optimization problem of unable to obtain large data sets fundamentally.

  DOI: 10.25236/mmmce.2020.034

2. Convolution neural network

Convolution neural network is a kind of feedforward neural network which contains convolution calculation and has a deep structure. With the development of deep learning theory, convolutional neural network has been developed rapidly and has been applied to computer vision, natural language processing and other fields. Convolution neural network is derived from artificial neural network. The basic structure of convolution neural network is convolution layer and pooling layer. The input is the original image, and the output of the former layer is the input of the latter layer. Thus, the deep feature of the target is extracted, and the weights are updated by back-propagation algorithm. Convolution neural network defines a method to extract local features, which can effectively respond to specific local patterns.

The schematic diagram of convolution operation is shown in Figure 1. Assuming the image size is 5 ×5, 3 ×3 convolution kernel is used to traverse the whole image. If the stride is 1, then a feature image with a size of 3×3 can be obtained by convolution kernel operation, and so on. How many new feature images of convolution response can be obtained by the number of convolution kernels.

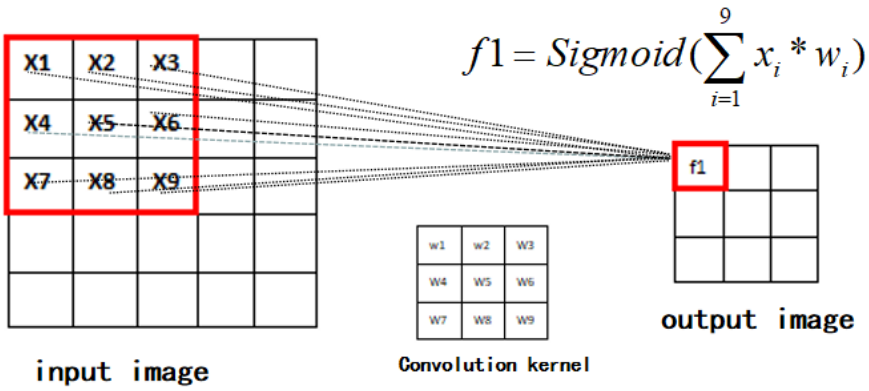

Figure 1. Schematic diagram of convolution principle

Pooling is also a common operation in convolutional neural networks. Pooling layer is a visual system that simulates human beings to reduce the dimension of data and represent images with higher level features. After convolution operation, there will be a lot of redundant pixel information in the output layer. Pooling operation is to reduce redundant information and improve the efficiency of feature extraction of convolution layer by fusing adjacent pixels. In addition, pooling can also improve the scale invariance of the model. Common pooling operations include maximum pooling, mean pooling, and median pooling. The process of pooling maximum values is shown in Figure 2.

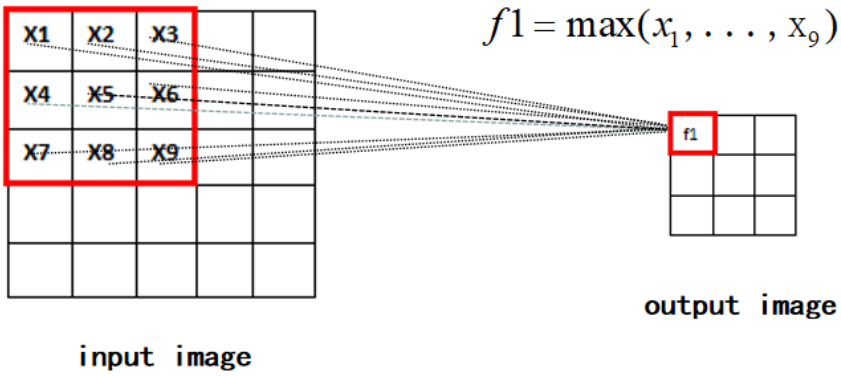

Figure 2. Schematic diagram of pooling principle

Convolution neural network is through a series of convolution layer and pooling layer accumulation layer by layer. After extracting the deep features of the image, the full connection layer is integrated to obtain the feature vector, and finally the classification layer is used for classification.

3. Acquisition and preprocessing of insulator dataset

3.1 Acquisition of insulator images

This paper starts from https://aistudio.baidu.com get the open source insulator dataset. The dataset includes pictures of 600 normal insulators and 248 defective insulators. The scale and state of insulator image are random, and the background of the image is complex, as shown in Figure 3.

(a) normal image (b)defective image

Figure 3. Normal/Defective images

3.2 Preprocessing of insulator images

Because the size, background and illumination intensity of images obtained from the website are quite different, image preprocessing is essential. Firstly, the vector median filter (VMF) is used to filter the color image, and then the image size is uniformly cut into 224 × 224, so as to meet the requirements of the classic AlexNet model for the image size. Finally, the self-defined dataset is labeled with the One-Hot Encoding.

4. Classic network structure

Deep convolution neural network originated from AlexNet network. In 2012, Alex et al. achieved the best results in the Imagenet competition image classification task by using the convolutional neural network, which made the convolutional neural network a great success in image classification.At present, most of the novel networks are proposed based on the classic networks such as AlexNet, VGG and GoogLeNet, and their structures are more complex than that of AlexNet. Therefore, it is more convenient for the rapid prediction and evaluation of the model to improve on the basis of AlexNet [4].

4.1 The network structure of AlexNet

The total number of layers in the AlexNet is 8, 5 convolution and 3 full connected layers. The first layer: convolution layer 1, input 224 × 224 × 3 image, the number of convolution kernel is 96, the size of convolution kernel is 11 × 11 × 3, stride = 4; Then proceed to (local response normalized), followed by pooling, pool_ size = (3,3), stride = 2, padding = 0, finally get the feature map of the first convolution layer.

Convolution layer 2, the input is the feature map of the previous convolution, and the number of convolution kernels is 256. The size of convolution kernel is: 5 × 5 × 48, padding = 2, stride = 1; then LRN is done, and finally max_ pooling, pool_ size = (3, 3), stride = 2.

According to figure 4, We can push out the rest of the class, but we need to improve the last full connected layer, because the number of classifications of ImageNet is 1000, while our insulator dataset has only two classification results, So the final output corresponds to two labels.

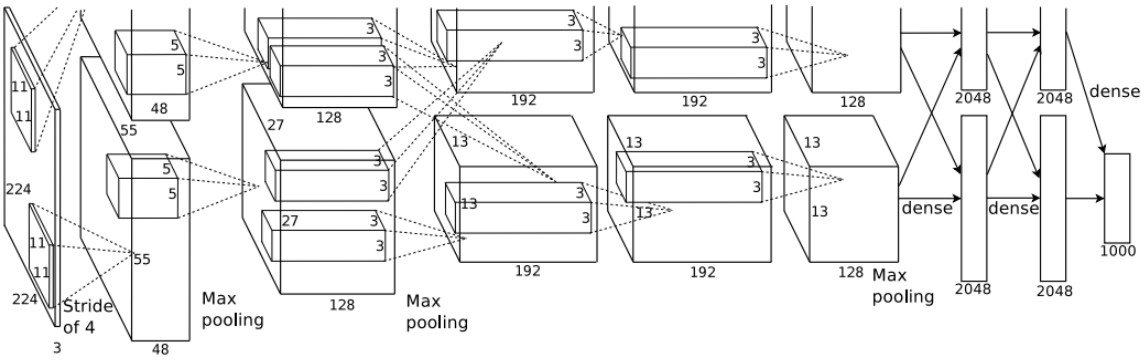

Figure 4. AlexNet network structure

5. Transfer Learning

Data mining and machine learning technology have achieved great success in classification, regression, clustering and other fields. However, many machine learning methods are carried out under the following assumptions: training data and test data are from the same distribution and the same feature space [5]. When the distribution changes, most machine learning methods need to recapture a large number of training data. However, in practical application, it takes a lot of time and cost to retrieve enough training data. Therefore, knowledge transfer or transfer learning between different fields emerges as the times require. Transfer learning has also become a popular method in the field of computer vision. It allows us to build accurate models in a time-saving way. In computer vision, we usually use pre-trained models to achieve transfer learning. The commonly used pre-training models are VGG, Inception and ResNet.

5.1 Strategies of transfer learning

(1) Generally speaking, there are two strategies for transfer learning:

fine tuning involves using a pre-training network on the base dataset and fine tuning all layers on the target dataset;

(2) Freeze and train, includes freezing and training only the last layer, keeping the other layers unchanged (weight is not updated), or freezing the previous layers to fine tune other layers. Specifically, it includes the following three options. As shown in Figure 5.

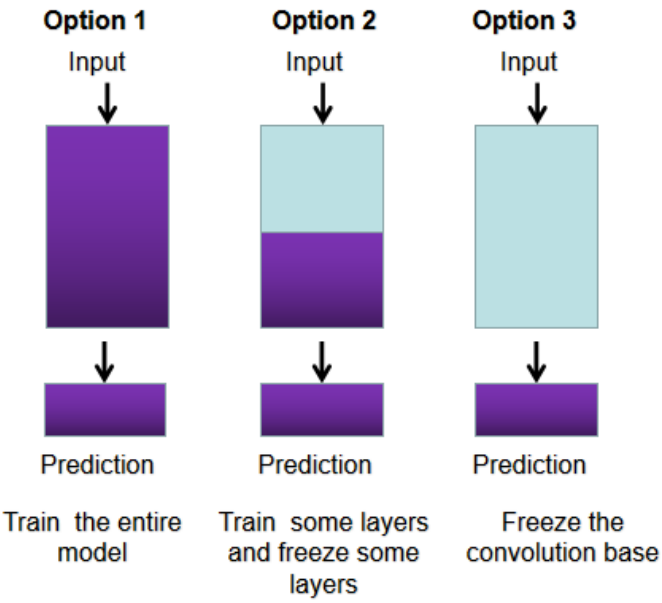

Figure 5. Three options of freeze and train

5.2 The realization process of transfer learning

The process of transfer learning includes: (1) selecting the pre training model;(2) The problem is classified according to the size similarity matrix;(3) Fine tuning model.

In this experiment, we choose the vgg16 modle built in keras in tensorflow 2.0. As show in figure 6, it includes five convolution groups and three full connected layers. The five convolution groups have 2, 2, 3, 3, 3 convolution layers respectively, so there are 16 layers.

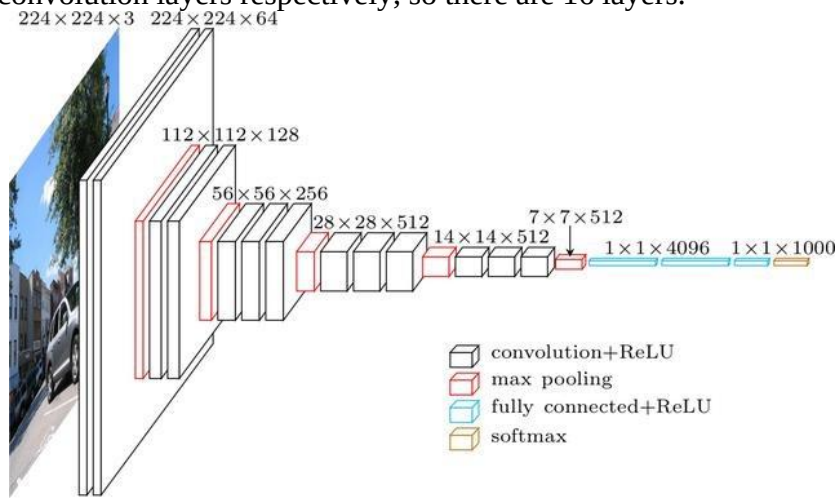

Figure 6. Vgg16 network structure

According to the size and similarity matrix shown in Figure 7, the insulator dataset we obtained is roughly in the D-Range. Therefore, in this experiment, we will try to remove the three full connection layers of the top layer of vgg16, retain most of the weights of the network structure and add a self-defined full connected layer to train the insulator dataset.The specific implementation process is shown in Figure 8.

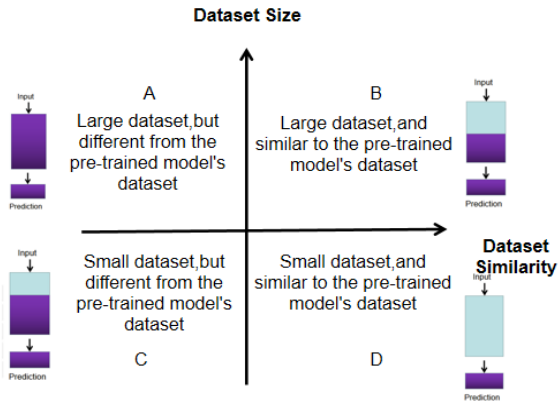

Figure 7. Dataset size and similarity matrix

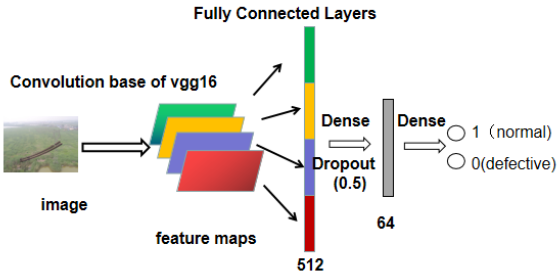

Figure 8. Fine-tuned vgg16 network structure

6. Experimental results and analysis

In the experiment, there are 848 images in the insulator defective detection dataset, including normal insulator and fault insulator. The size and location of fault area are different, and the background of fault area is various. 80% of them are training set for model training, and 10% are test set, which is used to test the accuracy of the current model after each epoch is completed; 10% is the verification set to verify the effectiveness of the trained model.

The final score results of the network are normalized to probability distribution by using softmax function. The calculation formula of softmax is:

$$S_i = \frac{e^i}{\sum_j e^j} \tag{1}$$

On this basis, crossentropy is used as the loss function to measure the difference between the two probability distributions, and the formula is as follows:

$$H_{(p,q)} = -\sum p_{(x)} \log q_{(x)} \tag{2}$$

p(x): the real probability distribution represented by the label
q (x): probability distribution of model represented by output

The optimizer selects stoichistic gradient descent (SGD), and its calculation formula is as follows:

$$\theta = \theta - \eta . \nabla_\theta J(\theta; x^{(i)}; y^{(i)}) \tag{3}$$

6.1 training and testing of models

The insulator fault detection method mainly includes two parts: 1) using the classic AlexNet network structure to realize the insulator fault detection and recording the convergence process of the network; 2) using the pre-trained model vgg16 after fine tuning to complete the insulator fault detection.

6.2 Analysis of experimental results

This experiment is based on tensorflow2.0, a deep learning framework. The batch of training set is set to 70 and the number of epoch is set to 10. Then the training starts. When the loss value of the model tends to be stable and the network converges, the optimal model of the network is solidified and various evaluation indexes are recorded, including crossentropy loss (loss), accuracy rate of the model on test set and validation set (acc/val_acc). The convergence process of AlexNet and the pre-trained network after fine tuning is shown in Figure 9.

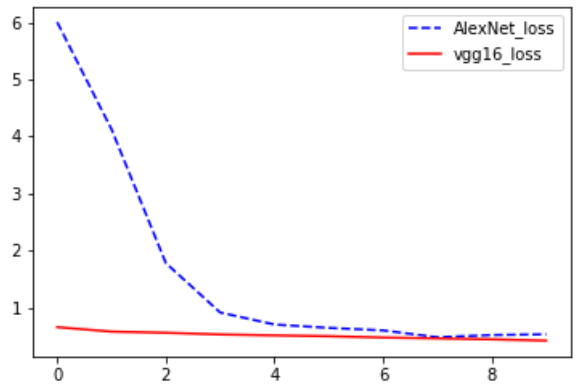

Figure 9. Convergence process of two models

It can be seen from the figure that the convergence speed of vgg16 is very fast, and it tends to be stable after the completion of the second epoch, while the AlexNet network gradually converges after several small amplitude oscillations. It can be seen that the reserved weight parameters of vgg16 network still have good effect on insulator dataset.

As shown in Table 1, after 10 epoch training, the accuracy rate of the pre-trained model after fine tuning has reached 82.53% on the test set and 88.05% on the verification set, while the classic AlexNet network performs poorly in both the test set and the verification set.

Table 1. Two models comparing

model	loss	acc	Val_acc
AlexNet	0.5308	0.7935	0.8333
Vgg16	0.4201	0.8253	0.8805

7. Conclusion

In this paper, firstly, the image detection principle of convolution neural network is proposed, and the alexnet network is set up and configured. After the training of custom insulator data set, the insulator fault detection model is obtained. Considering the randomness of insulator image state and the complexity of image background, the recognition accuracy of the model still reaches 83.33%. Compared with the traditional computer vision technology, it not only saves the tedious feature extraction process, but also performs well on the test set. In addition, in view of the small scale of the custom insulator data set, the model is in the state of under fitting, and the recognition accuracy cannot be improved. Based on the idea of migration learning, this paper introduces the vgg16 pre-trained network, freezes the weight parameters of the top layer of the vgg16 network, and adds the self-defined full connected layer. The new insulator fault detection model is trained to obtain the convergence speed and recognition accuracy effective improvement. It provides a certain reference value for power system and other fields to overcome the problem of dataset acquisition difficulty or small data scale in the future. However, due to the limited conditions, the insulator fault images used in this paper only contains self-exploding fault, and has not been identified for multi-class fault. Therefore, it is a problem to be solved further to expand the insulator fault image library and realize multi-category fault recognition.

References

[1] Tong Weiguo, Yuan Jinsha, Li Baoshu. Application of image processing technology in helicopter inspection of transmission lines [J]. Power grid technology, 2010,34 (12): 204-208

[2] Krizhevsky, Alex, Sutskever, Ilya, Hinton, Geoffrey E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM,2017,60 (6): 84－90.

[3] Wang Liming, Li Ang, Jianshe. Study on key factors of microwave nondestructive testing method for composite insulators [J]. High voltage technology, 2017,43 (01): 203-209.

[4] Shi Cuiping, Tan Cong, Zuo Jiang, Zhao Kexin. Facial expression recognition based on improved AlexNet convolutional neural network [J]. Telecommunication technology, 2020,60 (09): 1005-1012.

[5] Zhu Yingting. Markov retrieval model based on transfer learning theory [D]. Jiangxi Normal University, 2010