Multi-angle Insulator Recognition Method in Infrared Image Based on Parallel Deep Convolutional Neural Networks

Zhenbing Zhao[1(✉)], Xiaoqing Fan[1], Yincheng Qi[1], and Yongjie Zhai[2]

[1] School of Electrical and Electronic Engineering,
North China Electric Power University, Baoding 071003, China
`zhaozhenbing@ncepu.edu.cn`
[2] Department of Automation, North China Electric Power University,
Baoding 071003, China

Abstract. Insulator is the most common equipment in power system, and the recognition of insulator in infrared image is affected seriously under unconstrained condition with angle change. Due to the insulators' angle diversification and the manually-designed features' limitation, the existing algorithms have a problem of low accuracy in classification; therefore how to obtain the rotation invariant representations to cope with these adverse conditions is a very important problem. Deep Convolutional Neural Networks (DCNNs) have established a remarkable performance in image classification. The deep features obtained at the top fully-connected layer of the DCNNs (FC-features) exhibit rich global semantic information and are extremely effective in image classification. In this paper, we present a rotation invariant representation generation method named PFE-FDS (Parallel Feature Extraction and Feature Dimension Selection) for infrared insulator recognition which is based on parallel DCNNs FC-features extraction as well as feature sorting and dimension selection based on mutual information to eliminate redundancy. Then the SVM (Support Vector Machine) classifier is trained on our standard infrared insulator dataset for classification. The experimental accuracy shows that the PFE-FDS can improve the accuracy of multi-angle object recognition.

Keywords: Infrared insulators
Parallel deep convolutional neural networks
Rotation invariance representations · Feature dimension selection

1 Introduction

Insulator is the most common equipment in power system which is made of non-conducting material. It is used to support the electrical conductors and shield them from the ground or other conductors. The failure of insulators would be the direct threat to the stability and safety of the system [1]. With the advantages of

being non-contact and non-destructive, infrared imaging technology is efficient for monitoring and evaluating the thermal condition of insulators. According to statistics, tripping accidents caused by insulator fault accounted for the 81.3% of transmission line accident [2]. Therefore, monitoring insulator status regularly and detecting insulator fault timely is crucial. Accurate and efficient recognition of insulators is the premise of realizing the intelligent detection. Generally, insulators are diversified with different orientations, and using common feature representations for recognition may not be accurate, so obtaining robust rotation invariant representations is necessary.

The pictures in Fig. 1 show the positive and negative training samples as well as the test samples after multi-angle rotating.
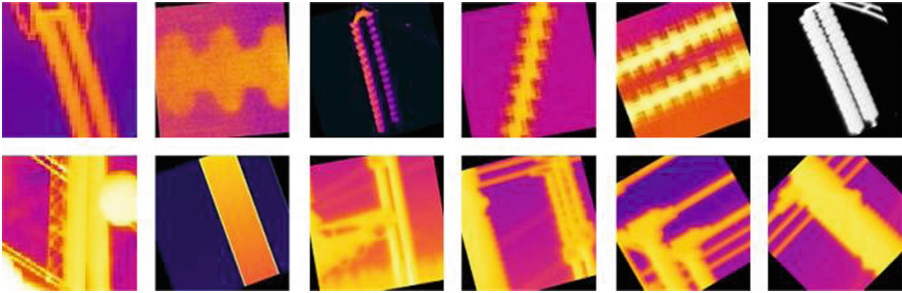

Fig. 1. Positive and negative samples. The two rows show the positive and negative images for training and testing with different orientations, respectively

Over the last few years, there has been some progress in insulators recognition. Yao et al. [3] proposed a zero value insulators recognition method under different pollution levels and different humidity conditions by combining the feature of insulator strings' relative temperature distribution characteristics and extracted from the artificial neural networks. Zhao et al. [4] extracted insulator outline from aerial insulator image based on non sampling contourlet transform. Jin et al. [5] extracted surface area and detected by using the optimal entropy threshold segmentation method. Ye [6] achieved object localization by using feature points matching between object image and template image based on SIFT (Scale Invariant Feature Transform) feature. Yen [7] studied the recognition and localization of insulators based on HOG (Histogram of Oriented Gradient) feature and SVM.

The above methods adopted the traditional hand-crafted feature and had a common problem of low accuracy, large calculation, being sensitive to the rotation of the angle. With the development of deep learning technology, more and more attention has been focused on the recognition of insulators based on the DCNNs. In this paper, we present a rotation invariant representation generation method for infrared insulator image named PFE-FDS which is based on parallel DCNNs FC-features extraction as well as feature sorting and dimension selection based on mutual information to eliminate redundancy.

2 Related Work

Recent researches show that DCNNs models can not only characterize large data variations but also learn a compact and discriminative feature representations when the size of the training data is sufficiently large, and it has good performance for object recognition and localization tasks [8]. Feature representations play an important role in computer vision, and have been widely used in many computer vision tasks [9]. An ideal feature representation should meet two basic characteristics, high quality representation and low computational cost, and it needs to capture important and unique information in images and be robust to various transformations.

After the AlexNet was proposed, more and more representative works have emerged in the structural optimization of the deep learning model. Zeiler and Fergus [10] get the new networks structure named ZF-net by reducing the size of the first convolution kernel in AlexNet from 11 * 11 to 7 * 7. Deep Convolutional Neural Networks have established an overwhelming presence in image classification starting with the 2012 ImageNet Large Scale Visual Recognition Challenge. In the VGG [11] model, 3 * 3 convolution kernels are used in convolution layers, which can significantly reduce the number of parameters and improve the discrimination. Residual Net (ResNet) [12] is a 152 layer networks invented by He et al., which was ten times deeper than others. Following the path VGG introduces, ResNet explores deeper structure with simple layer. Lin et al. [13] proposed a DeepBit32 model by introducing a new layer named fc8_kevin which encodes different representations of certain angle and obtains binary representations with certain rotation invariance.

DCNNs impose high computational burdens both at training and testing time, and training requires collecting and annotating large amounts of data. Recently, the second method has drawn much attention for image representation. Deep feature activations extracted from a pre-trained CNN model have been successfully utilized as general feature extractor for image representation. To get a generic representation, after a series of convolutional filtering and pooling, the neural activations from first or second fully connected layers (FC-layers) are extracted from a pre-trained CNN model. Gong et al. [14] proposed a certain scale pooling method to improve the rotation invariance of features. Yoo et al. [15] utilize Fisher Vector encode method for polymerizing multi-scale deep feature. The research of [16] shows that the ability of expressing features can be enhanced by integrating the representation of multiple layers. Tan et al. [17] proposed a Feature Generating Machine, which learns a binary indicator vector to show whether one dimension is useful or not. Deep representations generated from CNN model have achieved great success in object recognition. More feature selection methods [18–20] are proposed for different types of applications.

Current deep representations with small angle range rotation invariance can not meet the requirement of multi-angle insulators recognition. So this paper presents a rotation invariant representation generation method named PFE-FDS for infrared insulator image which is based on parallel DCNNs FC-features

extraction as well as feature sorting and dimension selection based on mutual information to eliminate redundancy. The details are described in Sect. 3.

3   Proposed Method

In this section, we propose a novel method named PFE-FDS to recognize insulators. Firstly, enter the input image into parallel DCNNs made up of pre-trained VGG16 model and DeepBit32 model, and extract different FC-layer feature representations. Then combine the representations directly and sort feature representations based on mutual information. After that select the dimension in line with the above sorting results. Finally the SVM classifier is trained on our standard infrared insulator dataset for classification. The overall framework is shown in Fig. 2.
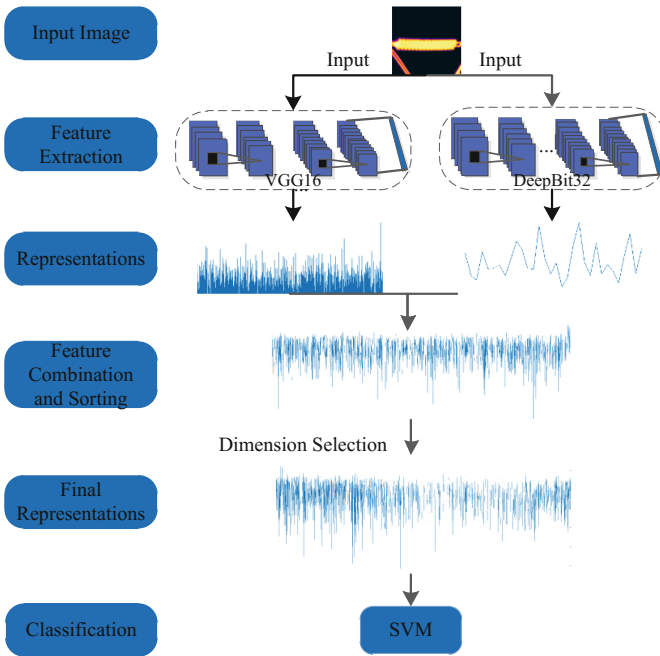

Fig. 2. Overall framework of PFE-FDS

3.1   Parallel Deep Feature Extraction

The research of [16] shows that the ability of expressing features can be enhanced by the integration of multiple layers. Inspired by this, our method extracts FC-features from parallel DCNNs. A typical DCNNs is made up of several convolutional layers, followed by pooling layers, fully-connected layers and a softmax decision layer.
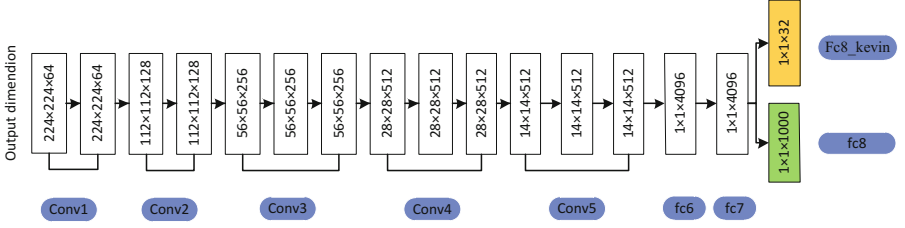
Fig. 3. Architecture of the DeepBit32 and VGG16 model trained on ImageNet 2012 classification dataset used in this paper. Each layer, represented by a box, is labeled with the size $R_l * C_l * K_l$ of its output in (1).

In Fig. 3 we illustrate the DCNN model. It consists of convolutional layers, max-pooling layers, fully connected layers and a softmax decision layer (It can also be called fc8 layer). At any given layer l, the layer's output data is an $R_l * C_l * K_l$ array

$$\left[x_{ij}^l \in R^{k_l}\right]_{i=1,\dots,R_l, j=1,\dots,C_l} \tag{1}$$

that is the input to the next layer, with the input to the first layer being an RGB image of size $R_0 * C_0$ and $K_0 = 3$ color channels. The fully connected layers can be seen as convolutional layers with kernels having the same size as the layer's input data.

While these methods adopted only the deep aspect of DCNNs, our goal is to combine the advantages of both approaches. The feature representations we utilize are extracted from fc6 layer in VGG16 model and fc8 layer in DeepBit32 model. The feature representations are complementary in discrimination and rotation.

3.2 Feature Combination and Sorting

In this section, we combine the features representations directly and form a representation of 4128 dimension. Because Principal Component Analysis (PCA) can not handle high-order correlation data, and can not be personalized optimization, so we utilize a feature selection algorithm based on mutual information for sorting. Although feature selection algorithm based on mutual information is not a new algorithm, it is the first application in the feature sorting of parallel DCNNs and power equipment recognition. Mutual Information is taken as the basic criterion to find Max-Relevance and Min-Redundancy between features [21]. The mutual information I of two variables x and y is defined based on their joint probabilistic distribution $p(x,y)$ and the respective marginal probabilities $p(x)$, $p(y)$:

$$I(x,y) = \sum_{i,j} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i) p(y_j)} \tag{2}$$

Max-Relevance is to search features satisfying (3), **S** denote the subset of features we are seeking, which approximates $D(S,c)$ with the mean value of all mutual

information values between feature x_i and class label c:

$$\max D\left(S,c\right), D = \frac{1}{S}\sum_{x_i \in S} I\left(x_i, c\right) \tag{3}$$

It is likely that features selected according to Max-Relevance could have rich redundancy. Therefore, the following Min-Redundancy condition can be added to select mutually exclusive features.

$$\min R\left(S\right), R = \frac{1}{S^2}\sum_{x_i, x_j \in S} I\left(x_i, x_j\right) \tag{4}$$

We define a new equation (5) for sorting. Then we use V_i denote the value calculated by (5) of the ith dimension feature, the values are arranged in descending order, and the ranking results are stored in the matrix $\mathbf{A}$ and the features are stored in $\mathbf{S}$ corresponds to the matrix $\mathbf{A}$.

$$V_i = \max\left(D\left(S,c\right)\right) - R\left(S\right) + \frac{D\left(S,c\right)}{R\left(S\right)}, i = 1, 2...4128 \tag{5}$$

3.3   Feature Dimension Selection

In image classification, the generated feature representations dimension is high and has rich redundancy. So we select the dimension of features $\mathbf{S}$ according to the classification results of SVM. The recognition result of VGG16_fc6 combine DeepBit32_fc8 is considered as a baseline. We select feature representations from matrix A in the top n for testing, the accuracy of recognition is the highest when n equals 3994.
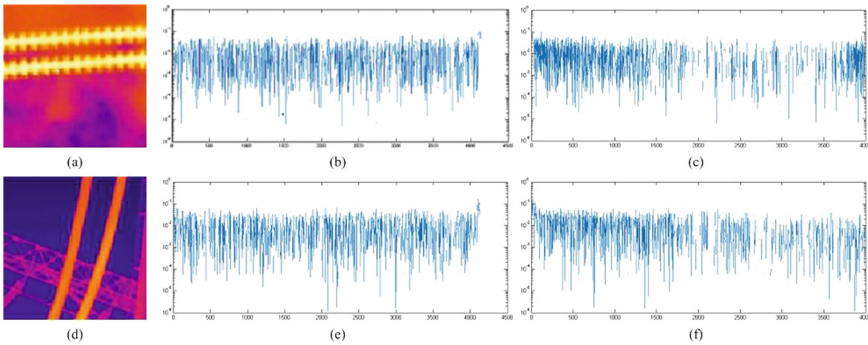

Fig. 4. Visualization of insulator and its feature representations

   As shown in Fig. 4, the picture (a) is the original insulator image, the picture (b) represents its feature representations of two model combined directly, the picture (c) represents its feature representations with dimension is 3994 after feature sorting and dimension selecting. The picture (d), (e), (f) is the feature representations of the corresponding negative sample.

4   Experiments Results and Analysis

In this section, we begin by introducing our infrared insulator datasets. Then, we evaluate our rotation invariant feature representations generation method on our standard infrared insulator datasets. In order to verify the practicability of this method, we selected two kinds of indoor scene for testing, and the effect is excellent.

4.1   Datasets

Due to the difficulty of obtaining insulators infrared image, and there is no public infrared image datasets, we use a large number of infrared images collected from insulator inspection system to build the insulator infrared image datasets. In the task of recognizing insulator, the infrared image datasets consists of 672 insulator samples and 1012 background samples. These original images are getting from the power substations varying from 110 kV to 500 kV level. Due to the limited samples of the insulator, we rotate the images manually for testing during the experiments.

We divide the dataset into two parts: 70% of this dataset for training and the remaining 30% for testing. All the training samples are labeled with "positive" and "negative", respectively.

4.2   Multi-angle Infrared Insulators Recognition

We visualize the feature maps of each convolutional layers of VGG16 model in Fig. 5. From the Fig. 5, we found that the neuron will response to the edge when the insulator image rotated and it will influence the recognition accuracy, but this influence is caused by human, can be ignored.
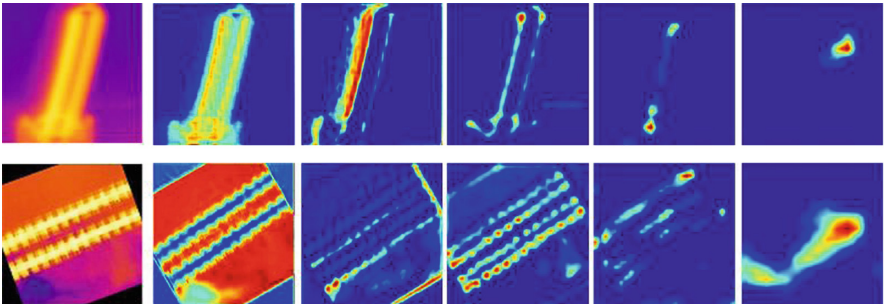

Fig. 5. Neural activation feature maps of each convolutional layers. The top line is normal insulator, the bottom line is the insulator rotated 30°

We extract feature representations from different FC-layers such as fc6 layer and fc8 layer in a DCNNs. Then we carry out the same operation on the different deep models such as VGG16, AlexNet, and we also experiment on some

traditional feature descriptors like Speeded Up Robust Features (SURF), Oriented FAST and Rotated BRIEF (ORB) and Binary Robust Invariant Scalable Keypoints (BRISK). Classification tasks are implemented on the normal samples and the samples rotating 30° respectively. The experimental results are shown in Fig. 6.
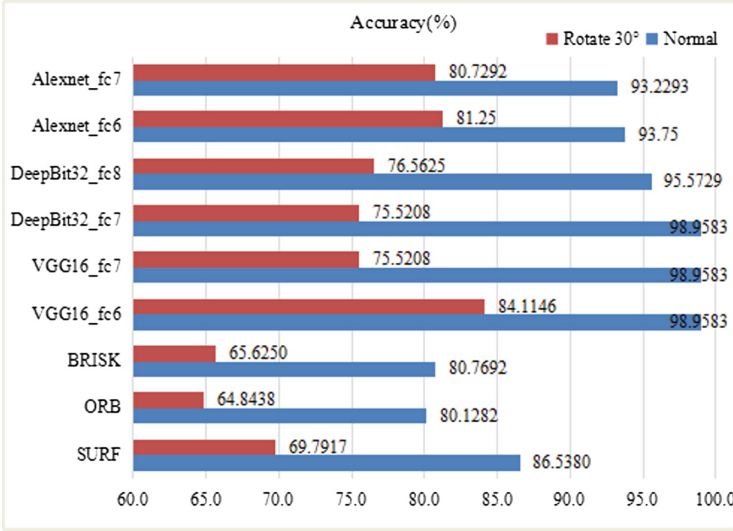

Fig. 6. Recognition results of different layers in different deep models (Color figure online)

Inspired by the results, whether it is to recognize the normal or rotating samples, the accuracy of VGG16 is the highest. Although DeepBit32 has rotation invariance in certain degree, its recognition performance is not good enough. And we discover the feature representations extract from VGG16 and DeepBit32 are complementary due to its introduced process. Based on this discovery, this paper will combine the two kinds of feature representations effectively, then sort features and select dimensions.

The test samples of datasets are rotated respectively and the rotation angle is 5°, 10°, 15°, 20°, 25°, 30°, 45°, 60°, then we utilize following representations for testing: (1) the feature representations from VGG16 fc6 layer; (2) the feature representations from DeepBit32 fc8 layer; (3) the feature representations combining two model representations directly, named P-DCNNs (parallel DCNNs); (4) the representations consist of two model representations, then conduct feature selection and no dimension selection, named PFE-FS (Parallel Feature Extraction and Feature Sorting); (5) the representations consist of two feature representations, then conduct feature selection and dimension selection which dimension is 3994, named PFE-FDS. The classification and intuitive results are shown in Table 1 and Fig. 7.
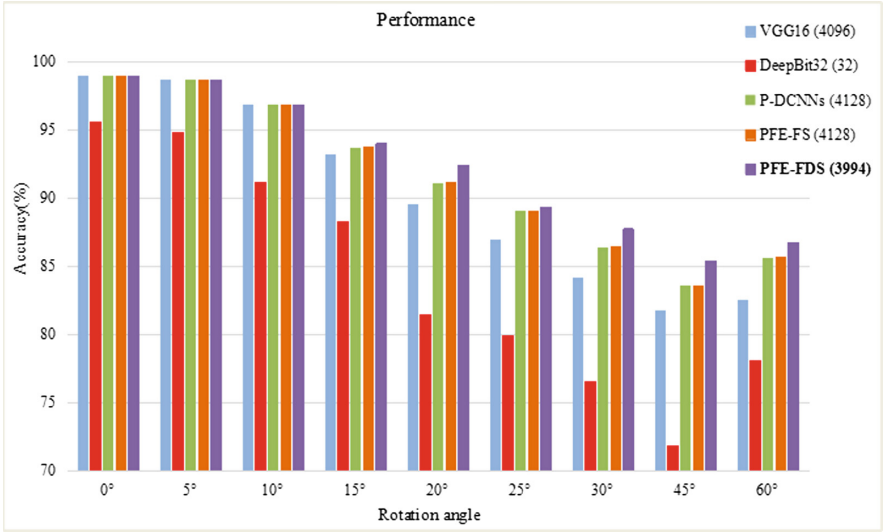
Fig. 7. Classification results of our method on multi-angle (Color figure online)

Table 1. Recognition results of insulators with multi-angle

Angles (°)	Accuracy (%)				
	VGG16 (4096)	DeepBit32 (32)	P-DCNNs (4128)	PFE-FS (4128)	PFE-FDS (3994)
0	98.9593	95.5729	98.9593	98.9593	98.9593
5	98.6979	94.7917	98.6979	98.6979	98.6979
10	96.8750	91.1667	96.8750	96.8750	96.8750
15	93.2292	88.2812	93.7500	93.7500	94.0104
20	89.5833	81.5104	91.1458	91.1458	92.4429
25	86.9792	79.9479	89.0625	89.0625	89.3229
30	84.1146	76.5625	86.4583	86.4583	87.7604
45	81.7708	71.8750	83.5938	83.5938	85.4167
60	82.5521	78.1250	85.6771	85.6771	86.7188

For the normal infrared insulators, the recognition accuracy of a single deep model has been very high, so the use of PFE-FDS is not much improvement, but with the angle increases, the recognition accuracy increases faster.

4.3 Multi-angle Scene Recognition

In order to verify the validity of the method we proposed, we select two kinds of similar indoor scene named airport inside and bar from the datasets published for the task of indoor scene recognition. The number of samples is 608 and 603 respectively, we divide the dataset into two parts: 70% of this dataset for training and the remaining 30% for testing, and we performed the same operation on the test set like the above. The scene dataset and its feature map of each convolutional layer are shown in Fig. 8, and experimental results are shown in Table 2.
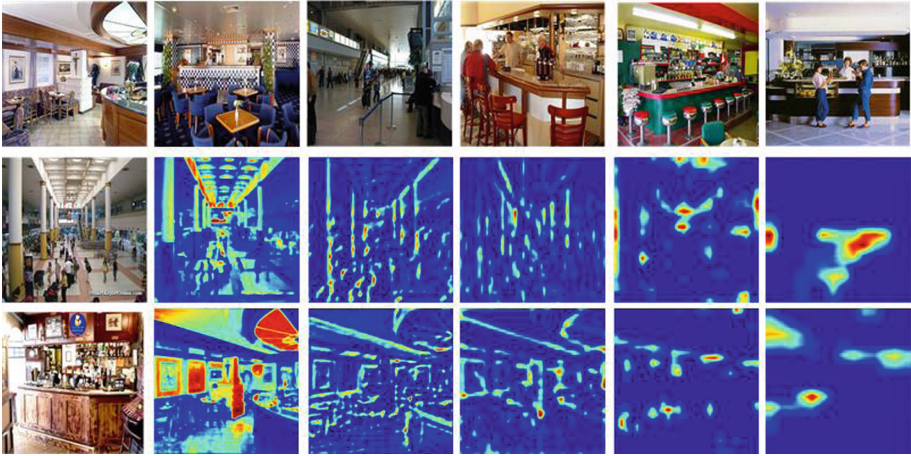
Fig. 8. Indoor image and its feature map

Table 2. Recognition results of indoor scene with multi-angle

Angles (°)	Accuracy (%)		
	P-DCNNs (4128)	PFE-FS (4128)	PFE-FDS (4004)
0	93.3649	93.3649	96.6825
5	92.8910	92.8910	95.2607
10	91.9431	91.9431	96.2085
15	94.3128	94.3128	96.6825
20	93.3649	93.3649	93.3649
25	92.4171	92.4171	92.4171
30	90.0474	90.0474	95.2607
45	83.4123	83.4123	85.7820
60	85.7820	85.7820	88.6256

For indoor scene recognition task, the recognition accuracy is highest when the dimension is 4004. The recognition results indicate that our proposed method outperforms the other two methods in precision.

From the results of the above experiments, we can see that the accuracy of our proposed method is not only higher than the two separate models, but higher than the accuracy of combining representations directly, and the contrast results of last two columns reflect the necessity select the feature and dimension. The experimental results of indoor scene recognition show that our proposed method is not only effective for multi-angle infrared insulator recognition, but also can be applied to multi-angle visible image recognition.

5   Conclusion

Infrared imaging has advantages in inspecting abnormal heating in electrical equipment, and it is efficient, reliable and non-destructive. However, most of the recognition and detection are conducted manually. With a great quantity of insulators to be inspected, the automatic recognition and localization method is needed.

In the light of the problem that the insulator recognition method is sensitive to the change of angles, and the recognition accuracy is low, a feature representation method for infrared insulators is proposed. Because of the few samples of infrared insulators, the data requirements of training model can not be met. We introduce the feature representations generation method PFE-FDS based on parallel DCNNs. The method doesn't need to do any direct finetune which needs a lot of time, realize the leap from feature designing to feature learning, then sort the feature and select its dimension to obtain the feature representations with robust rotation invariance and no redundancy. The high accuracy shows the efficiency of the proposed method. Then the method will be applied to practice.

References

1. Li, P., Huang, X., Zhao, L., Zhu, Y.: Research and design of transmission line condition monitoring agent. Proc. Csee **33**(16), 153–161 (2013)
2. Tong, W., Yuan, J., Li, B.: Application of image processing in patrol inspection of overhead transmission line by helicopter. Power Syst. Technol. **34**(12), 204–208 (2010)
3. Yao, J., Guan, S., Lu, J., Jiang, Z., Zhao, C., Xia, D., Qian, Y.: Identification of zero resistance insulators by combining relative temperature distribution characteristics with artificial neural network. Power Syst. Technol. **36**(2), 170–174 (2012)
4. Zhao, Z., Jin, S., Liu, Y.: Aerial insulator image edge extraction method based on NSCT. Chin. J. Sci. Instrum. **33**(9), 2045–2052 (2012)
5. Jin, L., Da, Z., Duan, S., Yao, S.: Contamination grades measurement of insulators based on image color feature fusion. J. Tongji Univ. **42**(10), 1611–1617 (2014)
6. Ye, H., Liu, Z.-G., Han, Z.-W., Yang, H.-M.: Fracture detection of ear pieces of catenary support devices of high-speed railway based on sift feature matching. J. Chin. Railw. Soc. **36**(2), 31–36 (2014)
7. Yan, L.: Insulator location and recognition algorithm based on hog characteristics and SVM. J. Transp. Eng. Inf. **13**(4), 53–60 (2015)
8. Zhang, T., Dong, Q., Hu, Z.: Pursuing face identity from view-specific representation to view-invariant representation. In: IEEE International Conference on Image Processing, pp. 3244–3248 (2016)
9. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: International Conference on Neural Information Processing Systems, pp. 1097–1105 (2012)

10. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8689, pp. 818–833. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10590-1_53

11. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. Computer Science (2014)

12. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition, pp. 770–778 (2015)

13. Lin, K., Lu, J., Chen, C.-S., Zhou, J.: Learning compact binary descriptors with unsupervised deep neural networks. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1183–1192 (2016)

14. Gong, Y., Wang, L., Guo, R., Lazebnik, S.: Multi-scale orderless pooling of deep convolutional activation features. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8695, pp. 392–407. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10584-0_26

15. Yoo, D., Park, S., Lee, J.-Y., Kweon, I.S.: Multi-scale pyramid pooling for deep convolutional representation. In: Computer Vision and Pattern Recognition Workshops, pp. 71–80 (2015)

16. Kulkarni, P., Zepeda, J., Jurie, F., Perez, P.: Hybrid multi-layer deep CNN/aggregator feature for image classification. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1379–1383 (2015)

17. Tan, M., Wang, L., Tsang, I.W.: Learning sparse SVM for feature selection on very high dimensional datasets. In: International Conference on Machine Learning, pp. 1047–1054 (2010)

18. Huang, K., Aviyente, S.: Wavelet feature selection for image classification. IEEE Trans. Image Process. Publ. IEEE Sign. Process. Soc. **17**(9), 1709 (2008)

19. Sun, Z., Wang, L., Tan, T.: Ordinal feature selection for iris and palmprint recognition. IEEE Trans. Image Process. Publ. IEEE Sign. Process. Soc. **23**(9), 3922–3934 (2014)

20. Zhang, Y., Wu, J., Cai, J.: Compact representation of high-dimensional feature vectors for large-scale image recognition and retrieval. IEEE Trans. Image Process. Publ. IEEE Sign. Process. Soc. **25**(5), 2407 (2016)

21. Mandal, M., Mukhopadhyay, A.: An improved minimum redundancy maximum relevance approach for feature selection in gene expression data. Procedia Technol. **10**(1), 20–27 (2013)