PURPOSE-LED
PUBLISHING™

PAPER • OPEN ACCESS

Study on Image Classification Algorithm Based on Improved DenseNet

View the article online for updates and enhancements.

Study on Image Classification Algorithm Based on Improved DenseNet

Gaihua Wang [a], Zhao Guo [b], Xizhou Wan [c], Xu Zheng [d]

Hubei University of Technology School of Electrical and Electronic Engineering Wuhan, China

[a] 20130006@hbut.edu.cn, [b] 201810237@hbut.edu.cn, [c] 1341516248@qq.com, [d] zhengxulion@163.com

Abstract. In order to perform fast analysis and recognition of power equipment images. A residual attention mechanism based image classification model is proposed. First, the power equipment dataset is preprocessed and the power equipment dataset is randomly divided into a training set and a test set. Based on the improved DenseNet, the image features of the training set are extracted. The training set images are used for training, and the classification test is implemented on the test set. The experimental results show that the accuracy of the improved DenseNet algorithm is improved by 8.89%. It can be seen that the classification accuracy of the improved DenseNet algorithm is effectively improved.

Keywords: Power equipment, DenseNet, residual attention mechanism, image classification.

1. Introduction

With the continuous development of convolutional neural networks, image classification technology [1] has been widely used in the fields of image search [2], target detection [3] and target localization [4]. Fast classification of power equipment images can not only solve the problem of low efficiency of manual long time detection [5] and classification [6] of power equipment, but also improve power equipment management and promote the development of power system. For images with complex backgrounds, it is still a difficult task to improve the classification accuracy of images.

The current image classification algorithms based on convolutional neural networks are mainly YOLO [7], VGG [8], ResNet [9], DenseNet [10], etc. Since the DenseNet network model is not applicable among any image classification, in order to improve the accuracy of the DenseNet network model in the classification of power equipment images, the corresponding improvements were made on the DenseNet network model [11].

2. DenseNet algorithm

DenseNet is a tightly connected convolutional neural network. In this network, there is a direct connection between any two layers. The input to each layer comes from all the previous output layers, and the feature maps learned by that layer will be transmitted as input directly to all the later layers.

The Densenet network model includes an input layer, a convolutional layer, a pooling layer, a dense block, a transition layer, and a softmax regression classifier. Convolutional neural networks usually need to reduce the feature map size by pooling or convolution in steps larger than 1, while the densely connected model of Densenet requires the same feature map size. To solve this problem, dense blocks and transition layers are used in Densenet. A dense block is a module with multiple layers, each of which has the same size feature map. The layers are tightly connected to each other. The transition layer module connects two adjacent dense blocks and reduces the size of the feature map by merging. the structure of DenseNet is shown in Figure 1, which contains four dense blocks. The reason for dividing DenseNet into multiple dense blocks is that the size of the feature maps within each dense block is desired to be uniform so that there is no size problem when making connections.
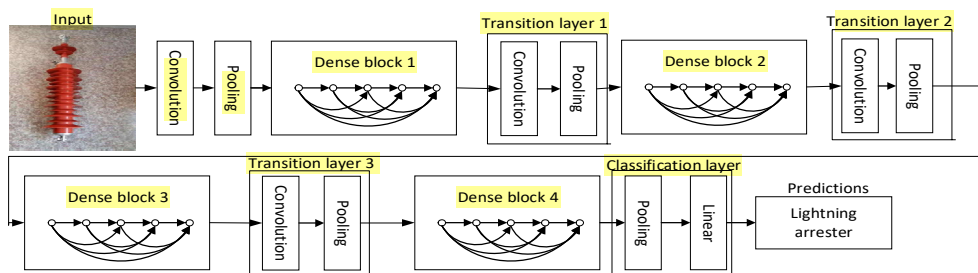

Fig.1 DenseNet network model

3. Improvement of DenseNet algorithm

In order to improve the classification accuracy of DenseNet network model on power equipment dataset, the superiority of residual attention mechanism can be used to enhance the effective features and suppress the invalid features. The residual attention mechanism is added to the DenseNet network model. The network model of the residual attention mechanism is shown in Figure 2.
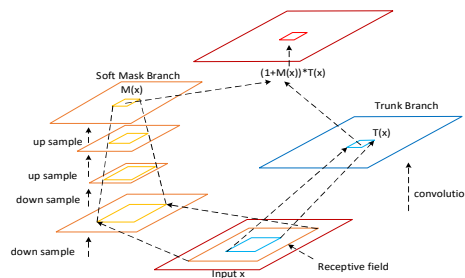

Fig.2 Network model of residual attention mechanism

In the figure above, x is the input, and the left is the mask branch, which first down-samples the network to allow the more significant information to be extracted, and then up-samples the output and input to become the same size. At this point, we get the attention map, which is represented as $M(\mathrm{x})$ in the top graph, and the trunk branch on the right side of this path, and the final output is also a feature map, which is represented as $T(\mathrm{x})$. Each value of each pixel in the attention graph output from the mask branch is equivalent to the weight of each pixel value on the original feature graph, which will enhance meaningful features and suppress nonsensical information. Therefore, the point multiplication of the feature map output from the mask branch and the trunk branch yields an Attention Mechanism network model output, whose expression is:

$$H_{i,c}(x) = \left(1 + M_{i,c}(x)\right) * T_{i,c}(x) \tag{1}$$

Eq. (1) represents the output of the Attention Mechanism network model, $M_{i,c}(\mathrm{x})$ represents the position of the ith pixel point, the output of the mask branch feature map corresponding to the position of the cth channel, and $T_{i,c}(\mathrm{x})$ represents the position of the ith pixel point, the output of the trunk branch feature map corresponding to the position of the cth channel, where the value of $M_{i,c}(\mathrm{x})$ ranges between 0 and 1, if the value of $M_{i,c}(\mathrm{x})$ is close to 0, then the value of $H_{i,c}(\mathrm{x})$ will be approximately equal to $T_{i,c}(\mathrm{x})$, so that in theory, the effect of having Attention Module will definitely not be worse than the effect of not having attention module. Here $M_{i,c}(\mathrm{x})$ is the key, by adding this mask branch, the model can adaptively enhance the significant features while neglecting the noise from the trunk branch.

DenseNet network is chosen as the base network, and a DenseNet network model based on residual attention mechanism is proposed for power equipment image classification on its basis. Since the residual attention mechanism can strengthen meaningful features and suppress meaningless features, after adding the residual attention to the DenseNet network model, the features will be better extracted and the feature representation will be stronger. After repeated experimental tests and comparisons, the best results were tested by adding the attention module of residual attention mechanism to the DenseNet network model between dense block 1 and transition layer 1; dense block 2 and transition layer 2; and dense block 3 and transition layer 3 simultaneously. The improved DenseNet network model contains input layer, convolutional layer, pooling layer, dense block, residual attention module, transition layer, and classification layer.The DenseNet-attention network model is shown in Figure 3.
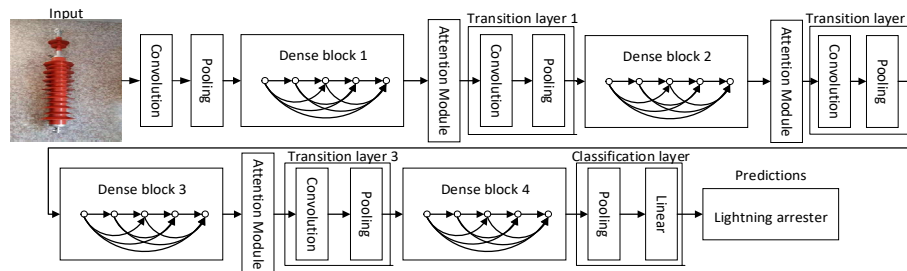

Fig.3 DenseNet-attention network model

4. Experiment and analysis

4.1. Introduction of the data set

A total of two datasets are used in this paper, the first dataset is the public dataset CIFAR-10, and the second dataset is the power equipment dataset.

The comparison experiments were first conducted on the public dataset CIFAR-10, which includes 10 different types of images: airplanes, cars, birds, cats, deer, dogs, frogs, horses, boats, and trucks. The dataset contains 60,000 images with a color image size of 32×32 size. The training set has a total of 5 files, each file contains 10,000 images, and the test set contains 10,000 images. The training set and test set images were previously divided. the CIFAR-10 dataset is shown in Figure 4.
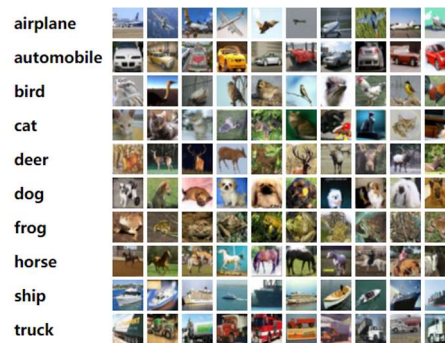
Fig.4 CIFAR-10 dataset

Finally, the test experiment is conducted on the power equipment dataset, which contains a total of three different types of power equipment. They are lightning arrester, insulator and transformer, respectively. Since there is no public and authoritative power equipment dataset on the Internet, the power equipment dataset needs to be produced by ourselves, and the images of the power equipment dataset in this paper are all obtained by self-organization. One part is obtained from the online Baidu download, and the other part is obtained from the field manual shooting of the target electric power equipment images to be classified. Each type of electric equipment images are 200, a total of 600 electric equipment images, because the obtained electric equipment dataset images are relatively small, so this paper expands the number of images in the dataset by data expansion methods, through mirroring, rotation and other operations, so that the number of each electric equipment image is expanded [12] to 6 times the original, so the total number of electric equipment dataset images is expanded to 3600, randomly 80% of the expanded images are selected as the training set and the remaining 20% as the test set. By expanding the images, the generalization ability of the classifier can be improved and the problem of overfitting caused by too few images in the dataset can be avoided. The images of the lightning arrester, insulator and transformer parts of the power equipment dataset are shown in Figure 5.
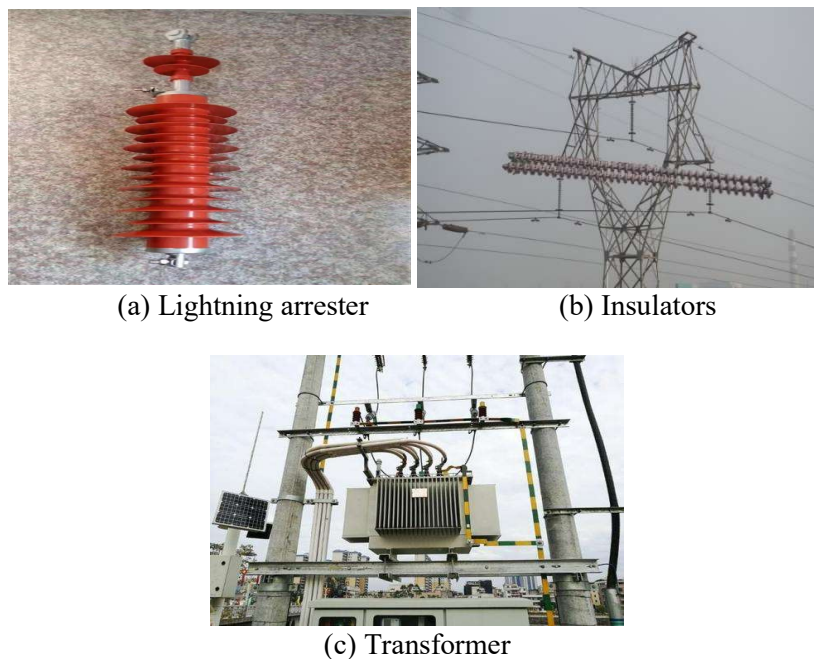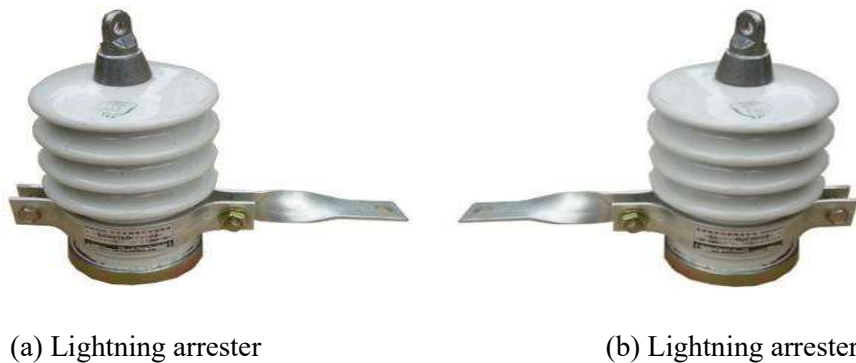

(a) Lightning arrester

(b) Insulators

(c) Transformer

**Fig.5** Picture of part of the power equipment dataset

4.2. Image widening

Image augmentation is performed by transforming the image to increase the number of images while ensuring that the image is not destroyed.

The following image augmentation techniques are commonly used: mirror transformation, rotation, scaling, cropping, panning, brightness modification, noise addition, clipping, and color change. Due to the limited experimental equipment conditions, this paper only adopts the rotate 90°, rotate 180° and mirror image techniques for image broadening, which increases the power equipment data set to 6 times of the original number of samples. The lightning arrester mirror image comparison diagram, transformer rotation 90° comparison diagram and insulator rotation 180° comparison diagram are shown in Figures 6, 7 and 8 respectively.

(a) Lightning arrester   (b) Lightning arrester

Fig.6 Lightning arrester mirror image comparison chart

(a) Transformer   (b) Transformer
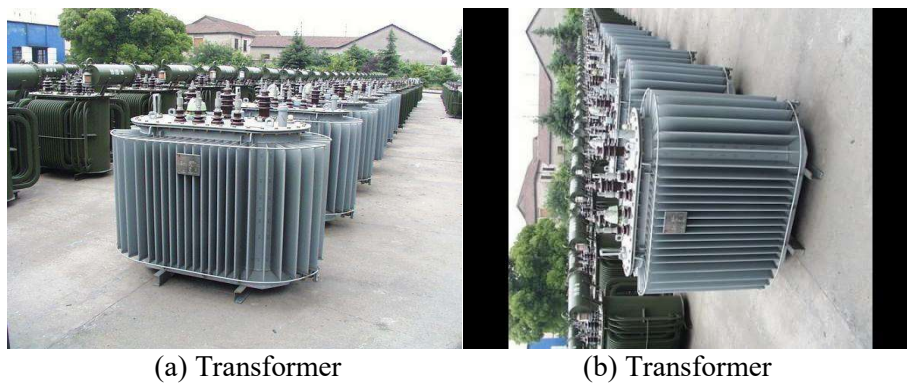
Fig.7 Transformer rotation 90° comparison chart

(a) Insulators   (b) Insulators
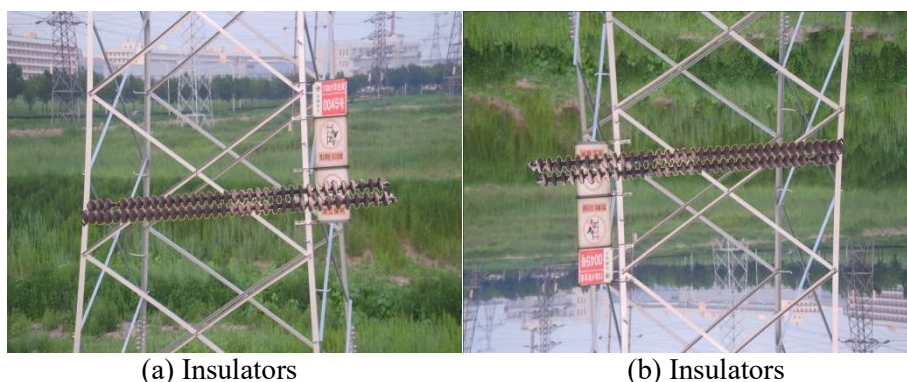
Fig.8 Insulator rotation 180° comparison chart

4.3. Comparative experiments

In this study, the experimental environment is Win10 operating system, the processor is Intel Xeon (Xeon) E5-2683 v3 @ 2.00GHz, and the running memory is 32G. In the experiment, the size of the public dataset is 32×32, the number of images per input model training is set to 32, and the number of iterations is 150 rounds. Since the size of power equipment images are different, firstly, the power equipment image dataset is preprocessed, and the image size is scaled to 224×224 uniformly in the program with the resize function. The number of iterations of the network model is set to 600 rounds, and the DenseNet network model is based on the Tensorflow framework, and the Tensorboard in Tensorflow is used to visualize the data obtained from the test. The accuracy and loss values of the test set of the improved DenseNet network and the original DenseNet network on the public data set and the power equipment data set are obtained by visualizing and entering the corresponding commands in the Anaconda prompt window, as shown in Figure 9.

(a) Public data sets

(b) Power equipment data set
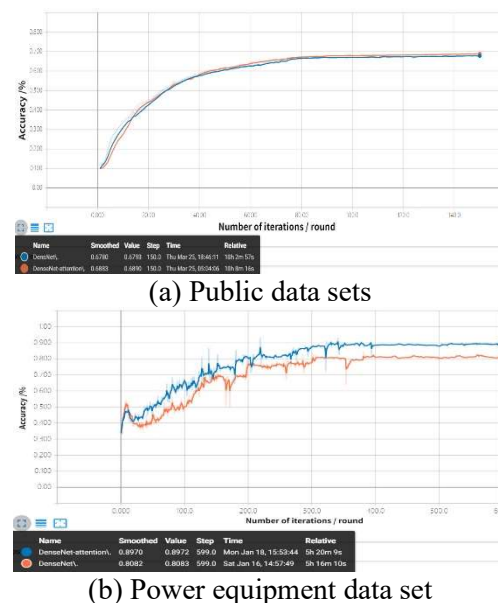
Fig.9 Test set accuracy

From Figure 9(a), it can be seen that on the public data set, the accuracy of the original DenseNet network model increases from 1 round, and the number of iterations to 80 rounds of accuracy does not change significantly. between 80 and 150 rounds, the accuracy basically remains stable. It means that the original DenseNet network model has reached stability, and the test accuracy is 67.93%. The accuracy of the improved DenseNet network model increases from 1 round to 80 rounds and basically reaches a stable state. The accuracy did not change much from 80 rounds to 150 rounds. The test accuracy was 68.90%. The improved DenseNet network model improved the test accuracy by 0.97% on the public data set.

As can be seen from Figure 9(b), the accuracy of the original DenseNet network model increases from 1 round on the power equipment dataset and remains basically stable after the number of iterations reaches 400 rounds. The test accuracy is 80.83%. The accuracy of the improved DenseNet network model increased from 1 round and remained stable after 400 iterations. The test accuracy was 89.72%. The accuracy of the improved DenseNet network model tested on the power equipment dataset improved by 8.89%.

The accuracy of the network model tested on the public data set power equipment dataset is shown in Table 1.

Table 1. Classification accuracy of network model

	Public data set CIFAR-10 (precision/%)	**Power equipment data set (precision/%)**
DenseNet	67.93	80.83
DenseNet-attention	68.90	89.72

5. Conclusion

In this paper, an improved DenseNet network structure is proposed based on the understanding of the DenseNet network structure. Adding a residual attention mechanism to DenseNet improves the effective extraction of image features by the network model. The improved DenseNet network is tested on the power equipment dataset than the original DenseNet network, and the testing accuracy is improved by 8.89%, although the time required for testing is increased. The application to power equipment image classification has some prospects. Because there is no publicly available power equipment dataset on the Internet, the power equipment dataset is a power equipment dataset made by downloading from the Internet and shooting on site. The types and numbers of power equipment acquired are limited. Therefore, the subsequent work will collect more types and quantities of power equipment datasets.

References

[1]   Shu Jun, Li Zhen-ya. Based on improved Inception-Resnet v2 image classification algorithm[J]. Journal of Hubei University of Technology, 2020, (5):62-64.

[2]   Huang Xiang-lin, Shen Lan-sun. Research on content-based image retrieval technology[J]. Journal of Electronics, 2002, 30(7):1065-1071.

[3]   Wan Ying, Han Yi, Lu Han-qing. Exploration of motion target detection algorithm[J]. Computer Simulation, 2006, 023(010):221-226.

[4]   Lei Zheng-guang. Research on the positioning of modern vocational education training objectives[J]. Vocational Education Forum, 2003, 000(009):6-9.

[5]   Bai Xue-peng, Liu Hua-pu. Application of RFID automatic identification technology in power equipment inspection information management [J]. Science and Technology Square, 2009, 000(003):134-136.

[6]   Wang Yu-ying, Zhao Qing-sheng, Liang Ding-kang. Deep learning network-based image classification of electrical equipment[J]. Science, Technology and Engineering, 2020, 20(23): 9491-9496.

[7]   Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger[C]// IEEE Conference on Computer Vision & Pattern Recognition. IEEE, 2017:6517-6525.

[8]   Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.

[9]   He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016.

[10]   Huang G, Liu Z, Laurens V, et al. Densely Connected Convolutional Networks[J]. IEEE Computer Society, 2016.

[11]   Fe I W, Jiang M, Chen Q, et al. Residual Attention Network for Image Classification[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017.

[12]   Wang Hai-wen, Qiu Xiao-hui. An Image Data Augmentation Method Based on Generative Adversarial Network[J]. Computer Technology and Development, 2020, 30(03):51-56.