Texture Recognition and Classification Based on Deep Learning

1st Gaoming Zhu
Collaborative Innovation Center for Modern Grain Circulation and Safety
Jiangsu Key Laboratory of Modern Logistics
College of Information Engineering
Nanjing University of Finance & Economics
Nanjing, China
1348525753@qq.com

2st Bingchan Li
School of Electrical and Automation Engineering
Jiangsu Maritime Institute
Nanjing, China
876364957@qq.com

3st Shuai Hong
Collaborative Innovation Center for Modern Grain Circulation and Safety
Jiangsu Key Laboratory of Modern Logistics
College of Information Engineering
Nanjing University of Finance & Economics
Nanjing, China
hjmrezl@outlook.com

4st Bo Mao*
Collaborative Innovation Center for Modern Grain Circulation and Safety
Jiangsu Key Laboratory of Modern Logistics
College of Information Engineering
Nanjing University of Finance & Economics
Nanjing, China
maoboo@gmail.com

Abstract—**Texture image classification has always been a very active research topic in computer vision and pattern recognition. In this paper, based on the deep learning advanced framework--Keras, we use Convolutional Neural Networks (CNN) to classify 12 kinds of texture images. Because there are few original datasets and the quantity is not balanced, We used such as reflection enhancement, elastic transformation, random lighting and other data *augmentation* techniques to enhance and expand some texture images. On the one hand, it balances the number of various types of texture images. On the other hand, it enhances the generalization ability of the datasets. It plays a key role in the training of the model and improves the accuracy of the model. The final test accuracy is close to 90%, which is more advanced and convenient than the traditional texture image classification method, and the accuracy rate is higher.**

Keywords—Texture Image, Deep learning, CNN, Data augmentation

I. Introduction

With the development of network transmission and terminal equipment, images have become one of the main sources of information for people, and the image data has grown explosively. Massive image data poses a challenge to information processing technology. How to effectively understand and use image data, and how to obtain useful information from a large number of image libraries has become a challenge and valuable research. Currently, image classification as a key technology in image analysis has been widely used in intelligent video surveillance, target tracking, content-based image retrieval and other aspects. Texture features, as one of the main underlying features of images, play an important role in image analysis.

We know that using texture features is an effective method when retrieving texture images with large differences such as thickness and density. However, when the difference between the texture information such as thickness and density is not large, It is difficult to accurately reflect the difference between different textures as seen by human vision. For example, reflections in water, the effects of reflections of smooth metal surfaces on each other all can cause texture changes. Since these are not the characteristics of the object itself, when texture information is applied to the search, sometimes the search will be "mislead" by these false textures.

In order to solve the "misleading" problem caused by these false textures, we use the convolutional neural network in deep learning to train a set of model for 12 class texture images recognition and classification. This model can accurately identify 12 common texture images and has strong practical application value.

II. Related Work

A. Deep learning and CNN

The concept of deep learning is proposed based on the research of artificial neural network. It aims to establish a mechanism which can simulate human brain to study the characteristics of data. Deep learning is a major breakthrough in the field of artificial intelligence in recent years. In 2006, Hinton put forward the concept of deep learning [1]. He proposed a multi-layer RBM structure based on probabilistic graph model -- DBN, which makes it possible to construct the artificial neural network and improves the learning ability of neural network. On the basis of Hinton's research, Ranzato [2] and Lee [3] et al use sparse coding mechanism to optimize the DBN. Tang and his colleagues [4] applied the noise reduction method to RBM, which further improved the learning performance of the artificial neural network.

In 2012, Professor Hinton (University of Toronto) and his student Krizhevsk won the ImageNet image classification champion by using GPU combined with deep learning. The classification accuracy of Hinton's team is 10% higher than the second. It is a very influential breakthrough for deep learning in the field of computer vision. In April 2013, MIT Technology Review ranked deep learning as one of the top 10 breakthrough technologies in 2013 [6].

Image recognition and classification is one of the earliest application fields of deep learning. Convolutional neural network (CNN) [7] is one of the mainstream deep learning models. It was proposed by LeCun and his colleagues in 1998 and was used for handwritten numbers recognition. Today, convolutional neural network (CNN) has been developed into an efficient image recognition method.

B. Texture Image Recognition and Classification

At present, most of the texture image recognition and classification are based on traditional feature extraction and classification methods. For example, in [8], based on the traditional LBP, an image texture feature representation method with anti-noise and rotation invariance is proposed. The article [9] uses SAM as a classifier and propose a texture image classification method that based on multi-feature extraction and constructing a look-up table to correct SAM classification results. Although these research methods have achieved good results to a certain extent, they are still relatively traditional in general, and they are actually more difficult to operate and have poor repeatability.

Now, This paper uses the Keras framework of deep learning to classify images. The trained model can be used repeatedly and gives more accurate predictions. The actual operation is more convenient and has stronger practical application value.

III. METHODOLOGY

In this paper, based on the deep learning advanced framework--Keras, we use Convolutional Neural Networks (CNN) to classify and classify 12 kinds of texture images. Due to the small number of original data sets and the unbalanced quantity, we have used reflection enhancement, elastic transformation, random lighting, and other data augmentation technology to enhance and expand some texture images. On the one hand, it balances the number of various types of texture images. On the other hand, it enhances the generalization ability of the datasets. It plays a key role in the training of the model and improves the accuracy of the model. The final test accuracy is close to 90%. which is more advanced and convenient than the traditional texture image classification method, and the accuracy rate is higher.

Experiment key points:

(1) The initialization of the parameter parser and some parameters: The image size of the datasets are not the same. Normalized dimensions are required during the training. After observation and analysis, I set norm_size=208, then, the input size of the image is (208,208).

(2) Load data: 12 types of pictures are marked with 0-11 tag, then images and corresponding tag information are loaded during training.The image and labels are shown in TABLE I.

TABLE I. IMAGE AND LABEL

name	beach	brick	building	cliff	tiles	crack
label	0	1	2	3	4	5
name	decal	door	planks	plaster	roofing	window
label	6	7	8	9	10	11

(3) We use the Adam optimizer. Since this task is a multi-classification problem, we use categorical_crossentropy as a loss function.

(4) Before and during the formal training, we used data augmentation technology to enhance our small dataset and enhance the generalization of the model.

A. Data Augmentation

The data used in this paper are 12 kinds of common texture pictures crawled on the network by means of crawlers. We use reflection enhancements and other data enhancement methods to enhance the images. The amount of data before and after data enhancement are shown in TABLE II and TABLE III.

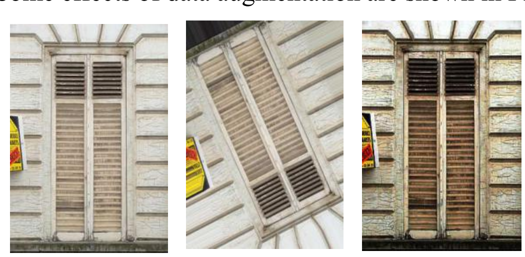
TABLE II. BEFORE DATA AUGMENTATION

name	beach	brick	building	cliff	tiles	crack
number	126	2163	2001	124	355	129
name	decal	door	planks	plaster	roofing	window
number	555	1991	672	1692	289	1045

TABLE III. AFTER DATA AUGMENTATION

name	beach	brick	building	cliff	tiles	crack
number	1940	2163	2001	1906	2243	2005
name	decal	door	planks	plaster	roofing	window
number	2081	1991	1903	1977	1903	2177

Some effects of data augmentation are shown in Fig1 :



(a) (b) (c)

Fig. 1. (a) original image; (b) the image after reflection and rotation transform; (c) the image after contrast stretch

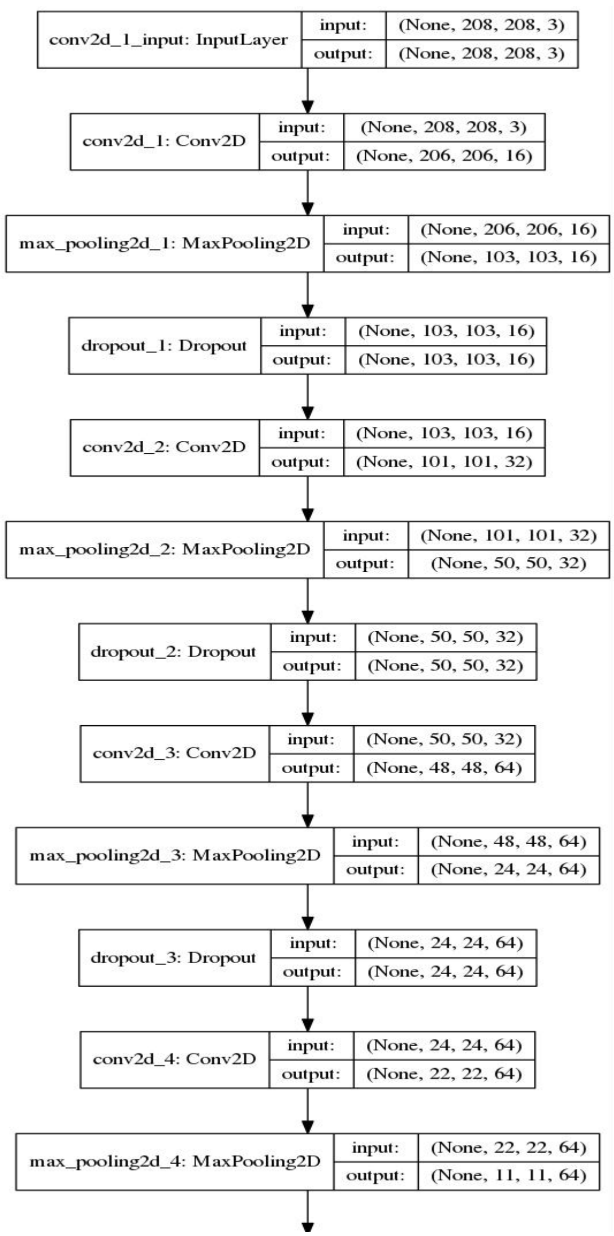
From the TABLE II and TABLE III, we can see that training datasets have been added. And from the Fig 1, we can see the effect of image after reflection and rotation transform and the effect of image after contrast stretch. Through these data augmentation methods, overfitting will be effectively avoided, and the accuracy of the model will be improved.

B. Convolutional neural network

In this paper, a specific CNN model for the current dataset is designed.

The input image is initially initialized to (208, 208, 3). This model has 5 convolution layers, 5 pooling layers, 6 Dropout layers, and 1 Flatten layer, 2 layers of fully connected image features, 1 layer of fully connected classification features.

The overall structure of the model is shown in the Fig 2 below.



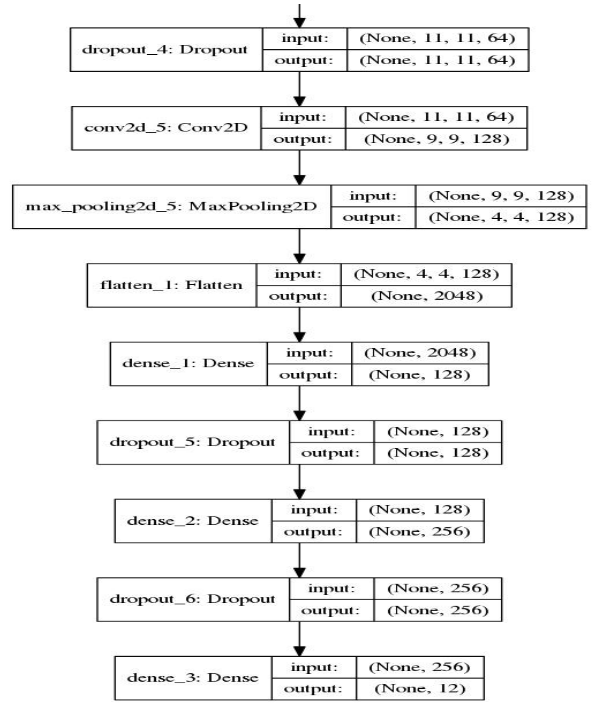


Fig. 2. the structure of the model

From the Fig 2, we can see the first layer is convoluted with 16 convolution kernels of size 3*3, the second layer convolved with 32 convolution kernels of size 3*3, and the third and fourth layers have 64 convolution kernels. There are 128 convolution kernels on the 5th convolution layer. The main role of the convolution kernel is to perform feature extraction and compression. Each convolution layer is added with "Relu" activation function. And each convolutional layer followed by a MaxPooling2D layer of size (2,2) and step size 2.Its role is to maximize the pooling of features exported by the previous layer. The six Dropout layers correspond to the regularization matrix. Dropout will randomly disconnect the input neurons at a probability of 0.2 each time the parameters are updated during training to prevent overfitting. In the transition from convolutional layer to fully connected layer, a Flatten layer is designed to "flatten" the input, the function of it is to make multidimensional input become one dimensional, and Flatten does not affect the size of the batch. The last layer of the fully connected layer uses softmax to classify 12 types of pictures.

IV. EXPERIMENT RESULTS

A. Two-class training and test

Before conducting training tests on all 12 types of texture images, we first took two types of beach and brick training tests. There are a total of about 4000 pictures of two types of image, and 75% are used as training datasets and 25% are used as test datasets. Set epoch to 10 and get the following result:

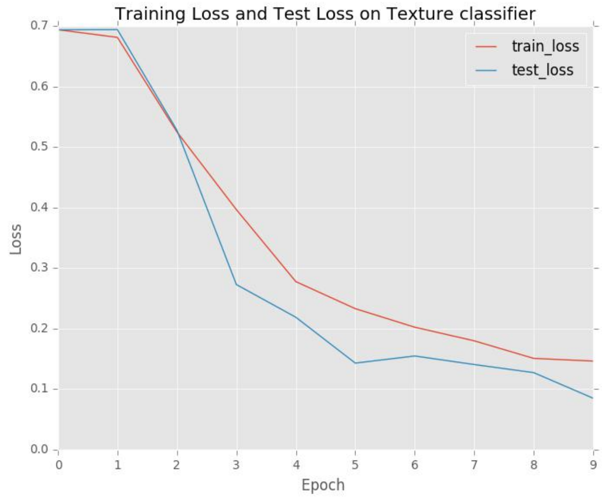


Fig. 3. the training loss and test loss in two-class

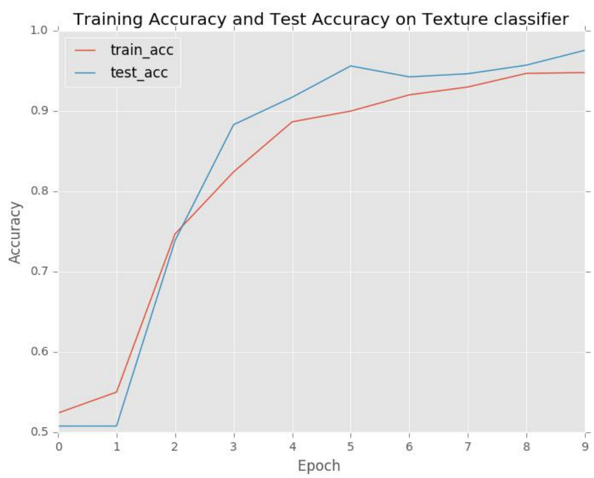


Fig. 4. the training acc and test acc in two-class

From the Fig3 and Fig4, we can see that after only 10 rounds of training, the test loss has dropped to 0.08, and the test accuracy rate has reached 0.98, which shows that our model is at least very effective for the two classification problems.

Next, we will test the accuracy of the model under multiple classification conditions (12 types of images).

B. Multi-class training and test

There are more than 20,000 images in the entire datasets, 75% of which are training datasets and 25% are test datasets. From Fig 5 below, It can be seen that after 100 epoch, the training loss and test loss decreased from more than 2 points at the beginning to about 0.3, and the overall loss showed a good downward trend..but at the same time it can be seen that the test loss dropped after 50 epoch become slow, and there are large fluctuations, but the overall result is normal, which may be related to the quality of our datasets, and may also need to be further optimized with our model. From Fig 6 below, we can see that the training accuracy has increased from the original 0.3 to the final 0.9, and the test accuracy has increased from the original 0.3 to the final 0.87. This result is more ideal. In [10], it also uses CNN to extract features, and has total of about 4452 images of 11 types of texture images, 75% of which are used for training datasets, 25% for test datasets, and use SVM to classification, and the final average accuracy rate is 79.78%. Compared with the article[10], we have a greate advantage. On the one hand, our total datasets and test datasets are five times of it. On the other hand, our accuracy rate is nearly 8 percentage points higher than it.

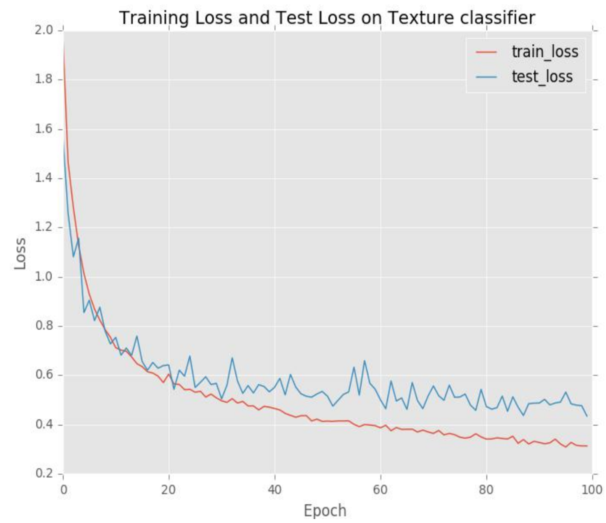


Fig. 5. the training loss and test loss in Multi-class

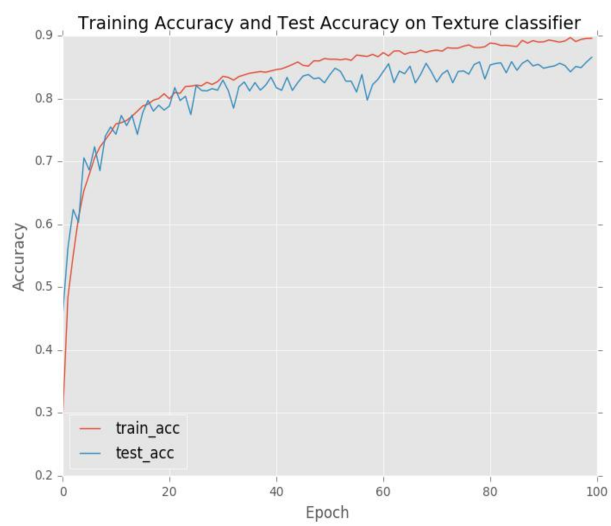


Fig. 6. the training acc and test acc in Multi-class

C. *Single image forecast*

After the model training is completed, we make a single picture prediction and get some prediction results as follows.

We use OpenCV to print the predict result and confidence on the picture.

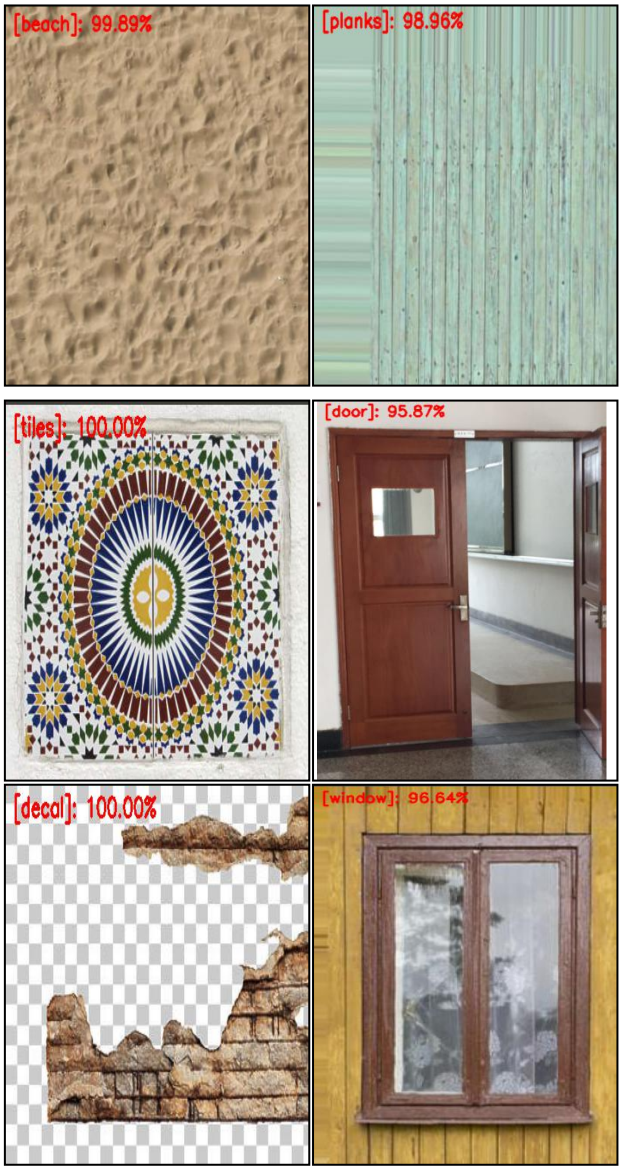


Fig. 7. the result of Single image forecast

From the Fig 7 above, we can see that every single picture has been predicted completely correct, and the confidence is close to 100%.

V. CONCLUSION

In this paper, based on the deep learning advanced framework--Keras, we use Convolutional Neural Networks (CNN) to classify 12 kinds of texture images. Because there are few original datasets and the quantity is not balanced, We used such as reflection enhancement, elastic transformation, random lighting and other data augmentation techniques to enhance and expand some texture images. On the one hand, it balances the number of various types of texture images. On the other hand, it enhances the generalization ability of the datasets. It plays a key role in the training of the model and improves the accuracy of the model. The final test accuracy is close to 90%, which is more advanced and convenient than the traditional texture image classification method, and the accuracy rate is higher.

More importantly, at the end of this paper, Single image forecast achieve greate recognition and prediction, which has a great practical value.

ACKNOWLEDGMENT

This work was supported by the National Key Technologies Research and Development Program of China under Grant (2015BAD18B02), National Science Foundation of China (41671457), China Special Fund for Grain-scientific Research in the Public Interest (201513004).Postgraduate Research & Practice Innovation Program of Jiangsu Province KYCX17_1211. Natural Science Foundation of the Higher Education Institutions of Jiangsu Province (16KJA170003, 14KJA520001). Collaborative Innovation Center for Modern Grain Circulation and Safety. Nanjing University of Finance and Economics.

REFERENCES

[1] Hinton, Geoffrey E, and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." science 313(5786), pp. 504-507, 2006

[2] Bureau, Y-Lan, and Yann L. Cun. "Sparse feature learning for deep belief networks." Advances in neural information processing systems, pp. 1185-1192, 2008.

[3] Lee, Honglak, Chaitanya Ekanadham, and Andrew Y. Ng. "Sparse deep belief net model for visual area V2." Advances in neural information processing systems, pp. 873-880, 2008.

[4] Tang, Yichuan, Ruslan Salakhutdinov, and Geoffrey Hinton. "Robust Boltzmann machines for recognition and denoising." Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, pp. 2264-2271, 2012.

[5] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems, pp. 1097-1105, 2012

[6] 10 Breakthrough Technologies 2013. MIT Technology Review, 2013-04-23.

[7] Le Cun, Yann, et al. "Gradient-based learning applied to document recognition." Proceedings of the IEEE 86.11 (1998): 2278-2324.

[8] Huang Chen, Fei Jiyou, Liu Xiaodong. Textural Feature Classification Method Based on Gaussian Local Binary Pattern[J].Application of Electronic Technique,2018,44(01):121-124

[9] Tang Yinfeng, Huang Zhiminget et al. Texture Image Classification Based on Multiple Feature Extraction and SVM Classifier[J].Computer Applications and Software,2011,28(06)

[10] Ji Zhong, Liu Qing, Nie Linhong, Pang Yanwei. Research on Texture Classification Based on Convolutional Neural Network[J]. Journal of Frontiers of Computer Science and Technology, 2016, 10(03): 389-397.