Design and implementation of face recognition system based on convolutional neural network

Yu Ma¹, HuaJun Wang*, Jun Wan¹, ZhenHeng Wang¹, Yao Yang¹, Cong Huang¹
¹Chengdu University of Technology, SiChuan Province, 610059, China
*Corresponding author’s e-mail: mayu6@stu.cdut.edu.cn

Abstract: The main features of face recognition are easy to use, quick to update, high flexibility, high accuracy, high environmental adaptability and so on. This passage of facial recognition technology based on convolutional neural network, which combines the present popular Python and Opencv Keras to identify development system, will make face recognition more efficient and more intelligent. And convolutional neural network based on face recognition data can improve accuracy and reduce the error rate through training. This experiment makes efforts to perfect the training model and enhance the influence of face recognition, so it could be more influential and smarter in higher precision applicable areas.

1. Introduction
The face feature analysis based on the convolutional nerve is more efficient than traditional image processing methods. Firstly, this method mixed together recent development of the deep learning technology and the features extracted. Then it combined various classifiers to recognize facial expressions, which was the unique feature of the image to conduct network training. Next the network training was started, which was based on the facial expression database of experimental data, including angry expressions, natural fear, disgust, panic, joy and sadness[1]. Like the previous image pretreatment, the effectiveness of face recognition in convolutional nerve was verified through the pretreatment, which was highly easy to use and easy to operate.

With economic development and effective improvement in various aspects, the pursuit of biometric identification technology is stronger, which makes it be in good development stage throughout these decades. The nature of human mind is self-stability, which is relatively strong, and there are individual differences. This physiological characteristic has become an ideal basis for experiments. Face recognition, compared with other recognition methods, can better reflect the characteristics directly and conveniently. The acceptance of users has increased correspondingly, so it has been far-reaching research and application.

2. Related work
Although China started late, the country has advanced considerably in the face recognition fields. Many research institutions and many IT companies have set up research groups based on face recognition technology, such as Shanghai Yinchen Intelligent Recognition Technology Co., LTD. Many universities in China have also set up research groups to develop facial recognition technology, including university of Electronic Science and Technology, National University of Defense Technology, Tsinghua University, Beijing Jiaotong University and so on. And all of these research groups have achieved not
only cognition technology, the realization of identity real-name system function and escort for the Olympic Games[2]. Making the opening and closing ceremonies of the Olympic Games able to identify tickets to personnel, has made a significant contribution to the Security system of the Olympic Games, providing decision-making basis. In addition, some institutions in society have set aside funds for research on facial recognition, including the National Science and Technology Support Program, the 863 Program and the Natural Science Foundation[3]. China's corresponding research in the accuracy and recognition speed has reached the high-profile level. Chinese face recognition technology is in a period of rapid development.

The research of this topic, based on convolutional neural network and face recognition system design, mainly includes three parts: to complete the collection of related image library, to complete the compilation of related module code, to complete the functional test of the written module. After face collection, the data will be loaded. And image could be trained to a more accurate degree through continuous training, then accurately retrieve the face part and complete face recognition.

3. The method

3.1 Convolutional neural networks
Convolutional neural network can more truly reflect the intrinsic correlation characteristics of data, compared with ordinary learning methods. The input layer, the convolutional layer, the sampling layer, the full connection layer and the output layer constitute the convolutional neural network model. Under normal circumstances, there are multiple convolutional layers and pooling layers, which are arranged alternately[4]. The next connection of the convolution layer must be the pooling layer, and similarly, the connection between the pooling layer must be the volume base layer. This is equivalent to the process of convolution, which is why convolutional neural networks are named. The basic structure is shown in Figure 1. Each convolution layer contains a set of characteristics to deal with more complex problems, which includes reducing weight sharing and pooling operation on the complicated model. So the multi-layer neural network, which contains nerve convolution, has better performance than traditional multi-layer neural network[5].

![Figure 1. Basic structure of neural network.](image)

3.2 Deep convolution layer and sub-sample layer
The layers involve the work of a deep convolutional neural network (D-CNN), which is scaled down by the convolution layer and sub-sample layers within a single layer. After this concept was known by Mamalet and Garcia.in and popularized by Simard, we change back to the posterior subsample layer and use two steps to convolve a single convolution layer. Patterns on images can be extracted by the following expression:
The $p_{t=e}$ and $p_{t}$ are input and output pattern mappings, $F$ is through the activation function that we use in our work. $m_{t}$ is the convolution kernel weight. $\theta_{t}$ represents offset, which represents the total number of input element mappings. $sq(t)q$ represents the small size of the horizontal convolution step, $sp(t)p$ represents the small size of the vertical convolution step, $Rq(t)$ and $Rp(t)$ are the width and height of the convolution kernel respectively. Among the $M(t-e)$, $A(t-e)$, the input height and width elements are mapped as follows:

$$A(t) = (A(t-c) - Rp(t))/sp(t) + 1$$  
$$M(t) = (M(t-c) - Rq(t))/sq(t) + 1$$

### 3.3. SGD algorithm

As the number of neural networks increases, the parameters increase several times. Updating parameters is a very important issue when train networks. In order to update the parameters of the supervised learning convolutional neural network, the algorithm is usually used. This experiment stochastic gradient descent was use because of its convergence rate. This also ensures that the parameters converge to the optimal solution, and one of the reasons is that its update speed is higher than that of the gradient descent algorithm.

$$f(x) = \sum_{i}^{n} f_{i}(w_{i}, x_{i}, y_{i})$$

$$w_{i+1} = w_{i} - \eta \sum_{i}^{n} \nabla f_{i}(w_{i}, x_{i}, y_{i})$$

In the first equation, $f$ is the activation function, $w$ is the weight, and $x_{i}, y_{i}$ are the input and output of layer $I$. The second equation is the updating of the parameter weights of the gradient descent method. The idea of stochastic gradient descent is to randomly select one in each medium $\nabla f_{i}$, not the one above $\sum \nabla f_{i}(w_{i}, x_{i}, y_{i})$, with the randomly selected direction as the direction of descent, as shown in the third equation. Figure 4-6 shows the feature activation calculation process in the algorithm.

### 4. Experiment and analysis

In this test, different faces and different poses were tested, including different angles. Therefore, I found that the test results changed slightly when we were in different poses. When we were facing the camera, the system captured the frontal photos and gained more information about faces, which enabled faster matching and more accurate matching results. When the face was tilted at a certain angle, the camera captured the face with the certain angle and could not harvest all facial features, so the speed of the test is affected. And when deviated to a certain angle, it also has a certain impact on the test results. The accuracy rate was almost 99 percent for frontal recognition, but for lateral recognition was only between 88 and 89 percent.

#### 4.1. Data preprocessing

The preprocessing mainly utilizes limited resources (computer video memory and data are limited) to optimize and improve the convolutional neural network model VGGNet. 16 and SqueezeNet and migrate them to face recognition tasks. First, this study analyzed basic structure of the VGGNet. 16 and SqueezeNet. According to the characteristics of the model and parameters, the experiment optimized parameters and the structure of network to meet the current hardware and data resources. Increasing the batch normalized layer made the network computing more smoothly, which prevented a fitting at the same time and improved the generalization ability. Increasing global average pooling layer, smoothing face feature and improving recognition rate were conducted before the network output. Then, experiments were carried out on the improved model. The face recognition rate was improved through the constant adjustment of structural parameters and hyperparameters of the model. Finally, the
experimental results and models were analyzed. Figure 2 shows the convolutional neural network training process.

4.2. Model training

Model training refers to optimizing the database, collecting information, updating and perfecting the database through constant repetition, so as to make the final test results more accurate and the process more rapid\[^4\]. In this area, scientists are not only looking for accuracy, but also for speed.

Amounts of trainings have been operated many times in this study. A lot of research has been done based on different optimizers. And several experiments have been conducted in different environments. SGD algorithm, Adam and Adadelta were used in these different environments, which made the final data more reliable and valuable. The experiments were also conducted during the day and at night, and all tests were correct. The simulated trainings of different optimizers were recorded, that obtained the change of the loss of training and the change of the accuracy of training. As shown in figure 3 below, the variation of training loss obtained for different optimizers in this study. And figure 4 shows the variation of training accuracy for different optimizers in this study.

After this training that can be learned, that the final recognition would be more accurate. As the more samples and faces were collected, and the database got larger. It can also be seen that SGD algorithm has the biggest change in training loss in the three different environments. Adam has the highest accuracy variation. Figure 5 shows the training process.
4.3. Experiment

The experiment of this test is completed in the environment of Python 3.5, OpenCV 3.3, Keras2.2.2 and TensorFlow 1.10.

As can be seen from figure 6, the accuracy of the network model designed in this paper and VGG16 is 97.35%. Currently, deep learning models are moving in a deeper and more complex direction. As the number of layers and levels increases, how to configure the network model without providing detailed theoretical support depends entirely on the task type and experimental results. The more complex the model is, the better the learning ability would get. Experimental results show that the deep convolutional neural network designed in this chapter can meet the requirements of the system.

| Model   | accuracy |
|---------|----------|
| VGG-6   | 97.35%   |

5. Conclusion

Convolutional neural network is used to build the model, mainly because the convolutional neural network is constant in geometric transformation, deformation, lighting, etc. And because of the sharing of the convolutional kernel, it takes less time to process high-dimensional data. A trained convolutional neural network can scan the entire image for detection at a low computational cost. The main contents are as follows:

(1) Analyze the current status of convolutional neural network in face recognition and explain the key issues of its research. Discuss research ideas.

(2) Construct the face recognition convolutional neural network model, that is, the deep neural network model composed of the convolution layer and the pooling layer alternately forms the first half of the model, and the second half of the model consists of multiple full connection layers and the last SoftMax layer. The random gradient descent algorithm is used to train the parameters in the network. At the same time, in order to prevent the over-fitting problem caused by the STOCHASTIC gradient descent algorithm, we add a "dropout" method to each layer of the network. This study mainly focused on face recognition, and made improvements in input data and training methods. During training, the method of supervision is canceled, so that the experiment can be carried out faster and the system can run more smoothly. The improvement of input data also makes the process run quickly and accelerates the implementation of the algorithm.

Reference

[1] Face recognition Face2vec based on deep learning: Small database case[J]. K. Sudars. Automatic Control and Computer Sciences. 2017

[2] Deep face recognition. Parkhi O M, Vedaldi A, Zisserman A. British Machine Vision Conference. 2015
[3] Facial expression feature extraction using hybrid PCA and LBP [J]. Yuan LUO, Cai WU, Yi ZHANG. The Journal of China Universities of Posts and Telecommunications. 2013 (2)

[4] ImageNet classification with deep convolutional neural networks. Krizhevsky A, Sutskever I, Hinton G.E. International Conference on Neural Information Processing Systems. 2012

[5] Eigenfaces for Recognition [J]. Matthew Turk, Alex Pentland. Journal of Cognitive Neuroscience. 1991 (1)

[6] Very deep convolutional networks for large-scale image recognition. SIMONYAN K, ZISSERMAN A. Computer science. 2014

[7] Jitender S D, Vijay V R, Hayri S. (2015) Data Mining: Research Trends, Challenges, and Applications. The Department of Computer Science University, Nebraska.

[8] Van der Geer, J., Hanraads, J.A.J., Lupton, R.A. (2010) The art of writing a scientific article. J. Sci. Commun., 163: 51–59.