Research on face recognition algorithm based on improved convolution neural network

Liu Hui

College of Information Science and Engineering

Wuhan University of Science and Technology

Wuhan,china
1711707819@qq.com

Song Yu-jie
College of Information Science and Engineering

Wuhan University of Science and Technology
Wuhan,china

447571628@qq.com

Abstract*—Convolution neural network is a depth learning algorithm that can automatically extract features. A face recognition method based on convolution neural network and Fisher criterion is brought up to resolve the difficulty of poor property of convolution neural networks under small samples. First,a discriminant metric function is added in the cost function of the error and enhance the classification of the network.Then the face features by utilizing the modified convolution neural network are extracted. Finally , the advantage of support vector machine in the small sample, nonlinear and high dimension is used to classify the extracted features. The experimental results show that the face recognition algorithm which based on the Fisher neural network combined with SVM can achieve good results in the case of fewer samples.

***Keywords*—convolution neural networks, Fisher criteria,face recognition,support vector machines,error cost function**

I. INTRODUCTION

Face recognition is a hot spot in the research of pattern recognition.It has a very wide range of applications in identification [1], credit card recognition [2] and monitoring system [3].However, facial images are also vulnerable to the external environment.Therefore, the extraction of facial features must take into account its stability and invariability.As a deep learning algorithm [4], convolution neural network has been widely used in image feature extraction because of the advantages of local receptivity [5] and shared weight.

There are many classifiers in face recognition which have been applied successfully.For example, Cheng[6] proposed a framework called Profile SVM (PSVM).It classifies the training samples into several classes and builds a linear SVM model on each class.Experiments show that the classification effect of this model is stronger than that of the global SVM classification.Softmax classifier[7] can deal with dichotomies and Multiple classification problem and it can calculate the probability of each sample to be tested. Therefore, it is widely used in neural network algorithms.

In recent years, deep learning has also been widely used in image recognition.Chen Yaodan, Wang Lianming and others put forward a 6-layer convolution neural network face recognition algorithm [8] (CNN).It trains weights and classifier parameters through a gradient descent method, which achieves good results in real-time and error rates.Tao Qinqin and Li Xiaohong proposed an algorithm for detecting face by combining convolution neural network with support vector machine [9] (SCNN).First,the The image features are automatically learned by using the convolution neural network, then the extracted features are to be classified by the SVM classifier.This method improves the performance of the convolution neural network, but the real-time performance of the algorithm is not high, and a large number of samples need training.Sun Yanfeng proposed a deep convolution neural identification algorithm based on improved Fisher criterion [10] (FCNN).It introduces Fisher criterion when training with deep convolution neural network, which makes the distance between non-same samples farther and the distance between similar samples closer. This algorithm can achieve good results in the case of small samples.

This paper presents a face recognition algorithm (FSCNN) based on improved deep convolution neural network and SVM.In the process of network training, this method adds Fisher discriminant criteria to the error function to control the distance between classes and classes.At the same time, in the process of classification of output layer, in order to improve the classification performance of the face images under the small sample, the features of the extraction were classified by support vector machine.The experiment results show that face recognition can achieve higher recognition rate when there are fewer samples.

II. CONVOLUTION NEURAL NETWORK AND CLASSIFIER

A. Convolution Neural Network

The convolution neural network is a network with multiple hierarchical network structures.It includes forward and reverse transmission.Forward transmission means that the input features go through multiple layers of operations and finally output features in the output layer.In each layer, an activation function needs to be added.Reverse transmission refers to make error calculation between the results of the forward transmission and the given sample tag firstly.And then transfer

978-1-5386-3758-6/18/$31.00 ©2018 IEEE

the error function back to each layer.Finally,the gradient descent method is used to update the network weights and bias parameters.

The error function that defines the *n*th sample expresses as follows.

$$E^n = \frac{1}{2}\sum_{k=1}^{C}(t_k^n - y_k^n) = \frac{1}{2}\| t^n - y^n \|_2^2 \tag{1}$$

Where t_k^n represents the *k*th label corresponding to the *n*th sample, y_k^n represents the *k*th output of the network corresponding the *n*th sample,*c* represents the dimension.

The output unit residuals on the lth layer expresses as follows.

$$\delta^l = \frac{\partial E^n}{\partial u^l} = f'(u^l)(y^n - t^n) \tag{2}$$

The partial derivatives of the error function for W^l and b^l express as follows:

$$\frac{\partial E^n}{\partial W^l} = \frac{\partial E^n}{\partial u^l}\frac{\partial u^l}{\partial W^l} = (\delta^l)^T x^{l-1} \tag{3}$$

$$\frac{\partial E^n}{\partial b^l} = \frac{\partial E^n}{\partial u^l}\frac{\partial u^l}{\partial b^l} = \delta^l \tag{4}$$

The lth layer iterative formula express as follows:

$$W^l = W^l - \alpha\frac{\partial E^n}{\partial W^l} \tag{5}$$

$$b^l = b^l - \alpha\frac{\partial E^n}{\partial b^l} \tag{6}$$

Where α represents the learning rate, and W^l and b^l represent the weight matrix and bias of the lth layer respectively.

B. *Convolution Neural Network with Discriminant Information*

In order to classify better, two kinds of metric functions named E_1 and E_2 are defined, which represent the distances of the same category and the distances between different categories respectively.

$$E_2 = \frac{1}{2}\sum_{c'=1}^{C}\| M_c - M_{c'} \|_2^2 \tag{7}$$

$$M_c = \frac{\sum_{n=1}^{N_c} y_c^n}{N_c} \tag{8}$$

Where y_c^n denotes the actual output value of the nth sample of class *c* ,and M_c and $M_{c'}$ denote the mean of the output of samples of type *c* and c' respectively, and N_c denotes the number of samples of type *c*.

The newly defined error function expresses as follows.

$$E = E^n + \lambda E_1 - \eta E_2 \tag{9}$$

Where λ and η represent weight parameters.

The newly defined layer *1* output unit residuals expresses as follows.

$$\begin{aligned}\delta'^{(l)} &= \frac{\partial E}{\partial u^l} = f'(u^l)(y^n - t^n)\\ &+\lambda(1-\frac{1}{N_c})f'(u^l)(y_c^n - M_c)\\ &-\eta\frac{1}{N_c}f'(u^l)\sum_{c'=1}^{C}(M_c - M_{c'})\end{aligned} \tag{10}$$

Then iterates and updates W^l and b^l according to the formula (5)and (6).

C. *Support Vector Machine Classifier*

Support Vector Machine (SVM) can transform the sample input space into a linear separable problem in high-dimensional space through nonlinear transformation when dealing with linearly inseparable problems.Then, the kernel function is used to replace the inner product operation in the high-dimensional feature space,thus find the optimal hyperplane for support vector machine (SVM).

The optimal discriminant function of support vector machine has the following form.

$$f(X) = \sum_{i=1}^{N} y_i a_i^* k(X_i \cdot X) + b^* \tag{11}$$

where $k(X_i \cdot X)$ represents inner product operation in feature space.

III. EXPERIMENTAL RESULTS

A. Experimental Data Set

This platform of this experiments.are Windows 7 and MATLAB 2014a.The experiments were carried out by using ORL face database and AR face database respectively. The ORL database contains 400 different face changes which are from 40 individuals and the AR database includes 2300 different face changes and 100 kinds of person.The convolution kernel has a convolution kernel size of 5x5.Down sampling layer using the maximum value of the delay, the sample size is 2x2.The activation function of the convolution layer uses the Sigmoid function.The network threshold is initialized to 0.SVM classifier chooses Gaussian function as kernel function..The learning rate α is 0.01.The weight parameters λ and η both take a value of 0.02.

B. Experimental Model

Figure 1 shows the system simulation structure.The model diagram includes SVM classifier, input layer, output layer, three convolution layers and two down sampling layers.

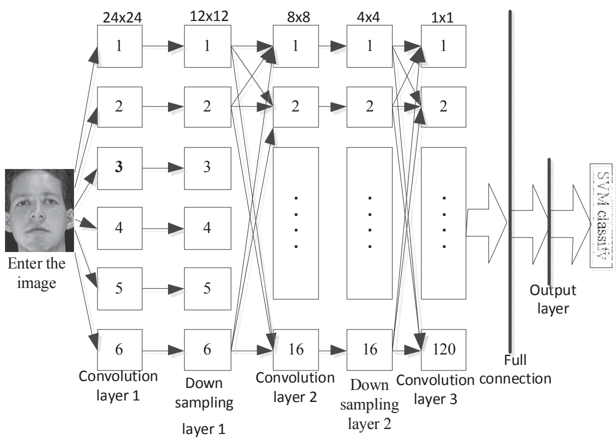

Fig. 1. System simulation of the structure

C. Experimental Results and Analysis

Several different methods were used to do comparative experiments for the ORL database.Taking 2, 4, 6 and 8 images from 10 images of each type for training, and other images for testing.CNN is a classical convolution neural network algorithm, SCNN is the algorithm [14], FCNN is the algorithm [15], FSCNN is the proposed method in this paper.The experimental results are shown in Table 1 and Figure 2.

It can be concluded from Figure 2 that when the number of samples is small, the proposed method has more advantages in recognition rate than other methods.When the number of samples per class reaches 8, there is little difference between the recognition rates of several methods.The data in Table 1 shows the comparison between the method proposed and the traditional CNN.It can be drawn from the table that when training time is low, the increase of recognition rate is higher.

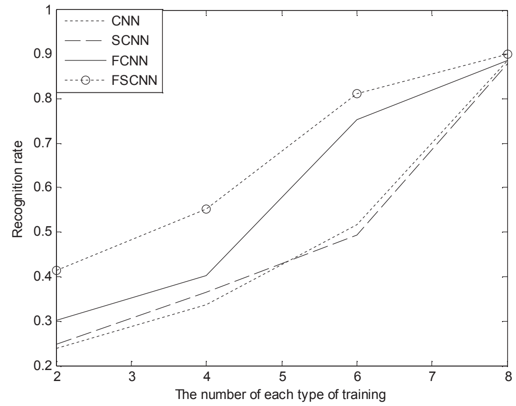

Fig. 2. Recognition rate of the different method when reducing training training sample in ORL database

TABLE I. COMPARISON OF RECOGNITION RATES BETWEEN CNN AND FSCNN WITH DIFFERENT TRAINING TIMES IN ORL DATABASE(%)

Training Times	CNN	FSCNN	Increase of Recognition Rate
300	68.32	80.35	12.03
500	78.54	89.72	11.18
1000	88.71	90.04	1.33

TABLE II. THE RECOGNITION RATE OF DIFFERENT METHODS WHEN THE NUMBER OF EACH TYPE OF TRAINING IS DIFFERENT IN AR FACE DATABASE(%)

Experimental Method	5	10	15	20
CNN	21.32	59.25	78.23	89.21
FCNN	28.45	63.29	81.35	88.34
SCNN	27.54	60.42	76.10	87.12
FSCNN	50.23	69.75	86.23	92.26

The same comparison test is also done for the AR database.For each category of 26 pictures,5,10,15 and 20 pictures were taken for training, and the rest are for testing.The experimental results are shown in Table 2:

The data in Table 2 shows that the recognition rate of several methods can reach about 89% with the increase of the number of samples, but the advantage of this method is relatively high when the number of samples is small.

IV. CONCLUSION

This paper presents a face recognition algorithm based on improved convolution neural network.The algorithm makes full use of the advantages of convolution neural network in feature extraction and the advantages of SVM in dealing with small samples and nonlinear problems.The proposed algorithm has certain reference value in some small sample recognition field.

REFERENCES

[1] Liyan Zhang,Anshuman Razdan,Gerald Farin,"3D face authentication and recognition based on bilateral symmetry analysis," The Visual Computer,vol.22, pp.43–55,2006.

[2] Feng Wang,Chin-Chen Chang,Wan-Li Lyu,"The credit card visual authentication schemebased on GF(28)field," Multimedia Toolsand Applications,vol.74,pp.11451–11465,2015.

[3] Syed Mohammad Ashraf,Ankur Gupta,Dinesh Kumar Choudhary,"Voltage stability monitoring of power systems using reduced networkand artificial neural network,"International Journal of ElectricalPower & Energy Systems,vol.87,pp.43–51,2017.

[4] Carlos Affonso,Andre Luis Debiaso Rossi,Fabio Henrique Antunes Vieira,"Deep learning for biological image classification,"Expert Systems with Applications,vol.85,pp.114–122,2017.

[5] Diana Turcsany,Andrzej Bargiela,Tomas Maul,"Local receptive field constrained deep networks,"Inform ation Sciences,vol.349,pp.229–247,2016.

[6] H Cheng,PN Tan,R Jin,"Localized Support Vector Machine and Its Efficient Algorithm,"In:SIAM Conference on Data Mining,pp.461–466,2007.

[7] Wang Hai-bo,Chen Yan-xiang,Li Yan-qiu,"Face Recognition Based on Principal Component Analysis and Softmax Regression Model,"Journal of Hefei University of Technology,vol.38,pp.759–763,2015.

[8] Chen Yao-dan,Wang Lian-ming,"Face recognition method based on convolutional neural network,"Journal of Northeastern University,vol.48,pp.70–76,2016.

[9] Tao Qin-qin,Li Xiao-hong,Zhan Shu,"Face detection based on convolutional neural network and improved support vector machine,"Hefei Polytechnic University,2016.

[10] Sun Yan-feng,Qi Guang-lei,Hu Yong-li,"Deep convolution neural network recognition algorithm based on improved Fisher criterion,"Journal of Beijing University of Technology,vol.41,pp.835–841,2015.