Face recognition algorithm based on Prewitt and convolutional neural network

Jiang Xiao
Shanghai University of Engineering Science
Shanghai , China
1225699783@qq.com

Gao Weiwei
Shanghai University of Engineering Science
Shanghai , China
gww03020234@sina.com

***Abstract*—In order to improve the recognition performance of face recognition algorithm, a convolutional neural network face recognition method based on Prewitt operator is proposed. Firstly, the face image is preprocessed by histogram equalization and Prewitt operator. Then, the preprocessed images are input into the convolutional neural network for training, and exponential decay method is adopted to set the learning rate to accelerate the convergence rate, and L2 regularization and Dropout are used to prevent overfitting. The recognition time of this method on ORL face database is 0.2s, and the recognition rate reaches 98.1%. The experimental results show that Prewitt operator and improved convolutional neural network can shorten the recognition time and improve the recognition rate.**

Keywords; Face recognition; Convolutional neural network ; Prewitt operator

I. INTRODUCTION

In recent years, face recognition has been widely used in human-computer interaction, access security system and other fields due to its high collectability, security and non-contact.With the development of technology, people have higher requirements on the accuracy of face recognition.

Traditional face recognition algorithms, such as improved principal component analysis (PCA) [1] proposed by DENG W, are used to describe the internal structure of a face. BELHUMEUR [2] proposed a face recognition algorithm based on LDA to reduce the overall dimension of face data and map it to low-dimensional space. These traditional face recognition methods are easily affected by lighting, expression, Angle and other factors, and require manual extraction of features, cumbersome process, low recognition rate and some limitations.

In recent years, the upsurge of deep learning initiated by HINTON[3] further promotes the development of convolutional neural network. Subsequently, the convolutional neural network, including GoogleNet, VGG[4],deepened the number of network layers and improved the recognition rate continuously. Although these deep convolutional neural networks can improve the face recognition rate continuously, the training difficulty and computation amount increase greatly with the deepening of layers. In order to solve the problems of low recognition rate of traditional face recognition algorithm, inconvenient feature extraction and large computation amount of deep network, a new method of improving convolutional neural network based on Prewitt operator is proposed.In this method, histogram equalization and Prewitt operator are used to simplify the facial features and reduce the computation after image compression. Exponential attenuation method is adopted to set the learning rate in the convolutional neural network. The activation function is Relu, and L2 regularization and Dropout are used to prevent overfitting.

978-1-7281-3584-7/19/$31.00 ©2019 IEEE

II. EXPERIMENTAL ENVIRONMENT AND DATA

The experimental hardware is configured with Intel Core i5-4200h CPU, 2.80ghz, 4G running memory, NVIDIA GEFORCE GTX 950M graphics card, Windows 10 system, python 3.7 programming language, and tensorflow 1.4 deep learning framework.

ORL face database is used as the data set for experimental training and testing.The ORL face library contains 40 people with different skin tones of different genders, each of whom has 10 photos with different expressions and details, and the rotation scale can be up to 20°. The background of all the photos is black, and the size is 112×92.

III. METHODS

Firstly, histogram equalization and Prewitt operator are used to extract edge features of human faces. Then, the extracted feature map is input into LeNet[5] convolutional neural network model, and exponential decay method, regularization and other algorithms are adopted to optimize the neural network and improve the recognition rate. The algorithm flow is shown in Fig. 1.

Authorized licensed use limited to: BEIHANG UNIVERSITY. Downloaded on June 15,2020 at 09:30:10 UTC from IEEE Xplore. Restrictions apply.

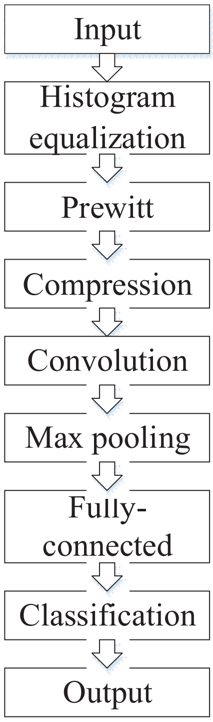

Figure 1 Algorithm flow chart

A. *image preprocessing*

1) face sample enhancement: The basic idea of histogram equalization is to transform the histogram of the sample image into a more uniform form, expand the dynamic range of pixel gray value , improve the image contrast, improve the gray value, and make the features of the sample more clear, which is conducive to the next edge feature extraction.

2) Prewitt operator: Prewitt operator is a kind of first-order differential operator, which is commonly used to detect image edges. Around the algorithm using a single pixel point and each point on the edge of the edge of the gray differential extremum to detect , in practical application are often used in the horizontal and vertical two direction of convolution to obtain corresponding direction in the field of template and the image gradient, but such rope rope rope if only use the vertical and horizontal directions template, the extracted edge characteristic information is not perfect, so by increasing the 45 ° and 135 ° two direction template, make facial features more complete. Prewitt operator can effectively suppress noise, and at the same time, it will not get false edges, and the image features after processing are more obvious.

B. *convolutional neural network*

Convolutional neural network (convolutional neural network) is a kind of incompletely connected multi-layer feedforward network, which has been widely used in the field of image processing because of its good fault tolerance, parallel processing ability and self-learning ability.Convolutional neural network consists of input layer, convolution layer, pooling layer, full connection layer and output layer.

1) Activation function: Sigmoid function or tanh function are generally selected by traditional convolutional neural networks when using activation function, but sigmoid function will appear gradient disappearance when training network [6], and the convergence speed is slow. As can be seen from Fig.2, the output of tanh function tends to be saturated with the continuous increase of input, so the Relu function is selected and the expression is as follows:

$$f(x) = \max(0, x) \tag{1}$$

This function has the advantages of unilateral suppression, relatively wide excitation boundary and sparse activation, making the generalization ability of the network stronger.

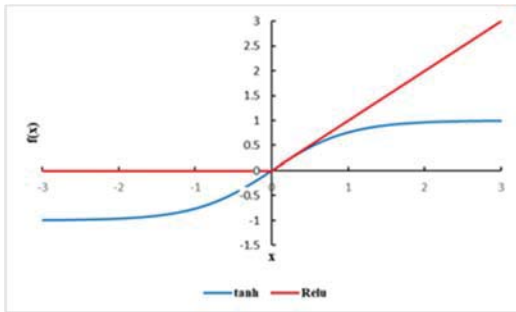

Fig.2 Tanh function and Relu function

2) exponential attenuation method: The updating range of parameters in training neural network depends on the learning rate.If the learning rate is set too large, the range of parameters will be too large each time they are updated, that is, the parameters fluctuate on both sides of the optimal solution and cannot be reached. If the learning rate is set too small, the convergence speed will be reduced. Therefore, exponential attenuation method is adopted to dynamically set the learning rate. The formula is as follows:

$$\eta = \eta_0 \varepsilon^{\frac{n}{\theta}} \tag{2}$$

In equation (2), η is the learning rate currently used; η_0 is the initial set learning rate; ε is the attenuation coefficient.After adopting this method, the network can quickly get a better solution by using a large learning rate. With the increase of the number of iterations, the learning rate gradually decreases, making the model tend to be stable in the later stage.

3) L2 regularization: The model can well fit each data or even noise on the known training data, while ignoring the overall trend of data, which is called overfitting.In order to prevent this situation, L2 regularization and Dropout technologies are adopted to ensure that the model after training has strong generalization ability.

L2 regularization is to add an index to describe the complexity of the model to the loss function, that is, to optimize the loss function $J(\theta)$ but $J(\theta) + \alpha R(w)$ during optimization. Where $R(w)$ represents the complexity of the model, and α represents the proportion of the complex losses of the model in all losses. Where E represents the weight in a neural network, and the calculation formula is as follows:

Authorized licensed use limited to: BEIHANG UNIVERSITY. Downloaded on June 15,2020 at 09:30:10 UTC from IEEE Xplore. Restrictions apply.

$$R(w) = \| w \|_2^2 = \sum_i | w_i^2 | \tag{3}$$

Since the full connection layer of convolutional neural network is prone to over-fitting, Dropout technology is adopted to randomly inactivate neurons in the full connection layer to prevent over-fitting.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The structural parameters of the convolutional neural network designed in this experiment are shown in table 1.

TABLEI. STRUCTURE PARAMETERS OF CONVOLUTIONAL NEURAL NETWORK

Layer	Patch size	Stride	Output size
Input	-	-	28×28×1
Convolution	3×3	1	28×28×16
Max pooling	2×2	2	14×14×16
Convolution	3×3	1	14×14×32
Max pooling	2×2	2	7×7×32
Fully-connected	-	-	512×1
Out put	-	-	40×1

Dynamic learning rate with continuous attenuation is used in training. Due to the number of iterations, the attenuation coefficient is generally set to a value close to 1, so it is set to 0.99 and Dropout value is set to 0.5.It can be seen from figure 3 that the speed of error rate decline in the training of the optimized network is significantly higher than that before optimization.

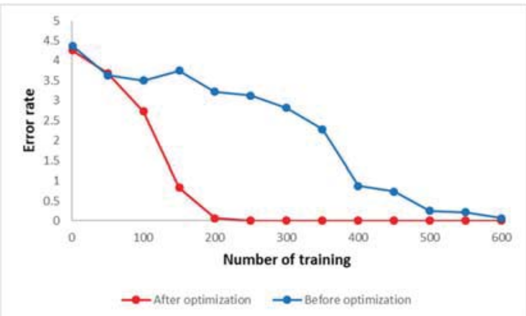

FIG. 3 comparison of decline rate of error rate

TABLE II. RECONGNITION RATE OF EACH FACE RECONGNITION ALGORITHM IN ORL FACE DATABASE

Methods	**Recognition rate /%**	**Time/s**
PCA	86.5	3.8
ICA	93.7	0.1
BP	81.5	5.2
CNN	94.4	0.3
proposed algorithm	98.1	0.2

To verify the effectiveness of this algorithm, PCA and ICA [7], two traditional face recognition algorithms, and BP and CNN, two algorithms that can automatically learn image features, are selected for comparison. Since traditional algorithms need to manually extract features and are easily affected by posture and occlusion, the recognition rate of Prewitt operator and convolutional neural network algorithm has been greatly improved, which is 11.6% higher than PCA, and the recognition time is 3.6s shorter.Compared with ICA, the recognition rate was improved by 4.4%.

Due to the large number of parameters involved in the operation and the tendency to fall into the local optimal solution, BP neural network has improved the recognition time and recognition rate by 17.2% compared with the following algorithm. Compared with traditional CNN, this algorithm USES histogram equalization and Prewitt operator in the image preprocessing stage, making the features of the image more clear and concise while reducing the parameters, so it improves the recognition time by 1s and the recognition rate by 3.7%.

V. CONCLUSION

Face image feature extraction is one of the keys to improve face recognition performance.Prewitt operator can effectively suppress noise without obtaining false edges. By preprocessing face images, features are simplified and strengthened, so that more differentiated information can be extracted in later neural networks.Experimental results on ORL face database show that the algorithm can effectively accelerate the convergence speed and improve the recognition rate, and has some advantages.

REFERENCES

[1] DENG W,HU J,LU J,.IEEE Transaction on Pattern Analysis & Machine Intelligence,2014,36(6):1275-1284.

[2] BELHUMEUR P N,HESPANHA J P,KRIEGMAN D J. “Pattern Analysis and Machine Intelligence”.IEEE Transactions on, 1997, Vol.1 (7), pp.711- 720.

[3] HITON G E, SALAKHUTDINOV R R . “Reducing the dimensionality of data with neural networks”.Science,2006,313(5786):504-507.

[4] YANG Z L ,YUE J.et al. “Vegetable Image Retrieval with Fine-tuning VGG Model and Image Hash”. IFAC PapersOnLine,2018,51(17).pp.280-285.

[5] LECUN Y,BOTTOU L,BENGIO P. “Gradient-based learning applied to document recognition” Proceedings of the IEEE,1998,86(11):2278-2323.

[6] HOCHREITER S,BENGIO Y,FRASCONI P “Gradient flow in recurrent nets: the diffuctly of learning long-term dependencies. Piscataway”. IEEE Press,2001:1-15.

[7] YANG M H. Kernel Eigenfaces vs. “ernel Fisherfaces: Face recognition using kernel methods”Nips,2002(2):1457-1464.

Authorized licensed use limited to: BEIHANG UNIVERSITY. Downloaded on June 15,2020 at 09:30:10 UTC from IEEE Xplore. Restrictions apply.