A Convolutional Neural Network Face Recognition Algorithm Based on Data Augmentation

Changyu Cao[1,a], Yang Wang[2,b], Chih Cheng Chen[3,c] and Jiachun Zheng[3,d]

[1]Navigation Institute Jimei University Xiamen, Fujian

[2]Navigation Institute Jimei University Xiamen, Fujian

[3]Information Engineering College Jimei University Xiamen, Fujian

[a]1103495664@qq.com, [b]1510109901@qq.com, [c]201761000018@jmu.edu.cn, [d]jchzheng@jmu.edu.cn

Abstract

Aiming at the problems of light changing and face blurring in face recognition based on convolution neural network,an optimization algorithm of data augmentationt is proposed, which uses LeNet and adds pool layer structure to extract face features.After the ORL database is enhanced, the training set and the test set are randomly generated for expeimental verification. The experimental results show that the accuracy of face recognition can reach 100% on the test set, and it has good robustness for some degree of illumination changes, facial expression changes and face blur.

Key words: Convolutional neural network, face recognition, data augmentation

Introduction

Face recognition is a key research object in the field of pattern recognition. It has broad application prospects in attendance, security, monitoring and other aspects. In recent years, with the rapid development and application of artificial intelligence [1][2] and the Internet of Things technology, face recognition technology has become an indispensable key technology of the current intelligent system. At the same time, it has also become the current research hotspot[3][4].

The research of face recognition began in the 1960s. The main achievements are as follows:the method based on geometric structure features was proposed by Bledson[5] et al, which consider the relative positions of key parts of face, such as nose tip, mouth angle as face features.The method is simple and less calculation, but the effect of illumination and expression change is general. R. Brunelli and T. Poggio[6] proposed a template matching method, which realized recognition by calculating the autocorrelation between the template and the image gray level, but neglected the local features, resulting in partial loss of information. M. Turk and A. Pentland proposed the method Eigenface [7]. Subspatial analysis algorithm Fisherface[8]. Based on artificial neural network [9], the method can extract features automatically, but the training time is too long. Face recognition method based on sparse representation [10][11] is simple to train, but it is time-consuming to solve sparse representation in practical application. In recent years, deep convolution neural network has successfully applied [12] in image classification, detection and other fields. Face recognition based on deep convolution neural network has opened up a new era of face recognition. In 2015, Google's team proposed a Facenet face recognition system based on Triplet Loss training, which was tested on LFW with 99.63% [13].Baidu research institute LFW recognition rate is 99.77%, tencent youtu reache 99.8%. Although these teams have a high recognition rate, they are based on large databases, large platforms for research, not suitable for individuals.In the research of face recognition loss function, Sphere Face normalizes the features to the projection sphere and designs a new loss, which also achieves good results [14].

In this paper, for the application scenarios of small platforms and small databases, the convolution neural network face recognition algorithm based on data augmentation (DACNN) is used to study the face recognition problem with different light changes and different blurriness of face images. The ORL database is enhanced by using image brightness, using image contrast and using image blur. At the end of training, face features are tested in the test set that generated by DAORL database(merged data augmentation with ORL data).

Convolutional Neural Network Model

The convolutional neural network of this paper has been adjusted on the LeNet structure, using data enhancement layer, three convolution layers, three pooling layers, two fully connected layers and one output layer. The structure of the DACNN neural network is shown in Fig.1.

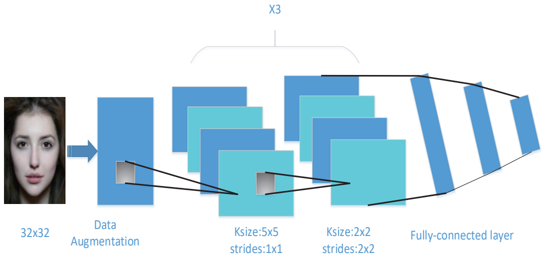

Fig.1 The structure of the DACNN neural network

A. *Convolution Layer*

The main function of the convolutional layer is to extract picture features, which have the characteristics of local perception and sharing weight. The convolution kernel parameter is 5x5 and has a step size of 1 in this paper. The convolution kernel is randomly initialized by truncated normal distribution function, and the bias initialized to zero. The mathematical expression is as follows:

$$x_j^l = f\left(\sum_{i\in M} x_i^{l-1} k_{ij}^l + b_j^l\right) \qquad （1）$$

doi: 10.35745/icice2018v2.037

l represents the current layer, k represents the convolution kernel weight, b represents the bias, and M_j represents the feature map.

B. *Pooling layer*

Pooling layer has two ways of maximum pooling and average pooling. it has the effect of reducing the feature graph dimension. it does not have the learning ability and does not change the number of feature graphs. In this paper, the maximum pooling method is adopted, the pooling size is 2x2, and the pooling step is 2.

C. *Fully Connected Layer*

The fully connected layer not only enhances the ability of network nonlinear mapping, but also limits the size of the network.The fully connected layers between the same layers are independent of each other and are not connected. The mathematical expression is as follows:

$$y_{pj} = f\left(\sum_{i=1}^{n} x_i^{l-1} w_{ji}^{l} + b_j^{l}\right) \qquad (\qquad 2@ \quad (2)$$

n represents the number of neurons in the upper layer, l represents the current number of layers, w_{ji}^{l} represents the connection weight between the layer of neurons j and the upper layer of neurons i, b_j^{l} represents the bias of the layer of neurons j.

Network Training

The ORL database has 400 pictures, each with 10 pictures and variations in facial expression, face size, and face posture. All images are 92x112 grayscale picture and the background color is black. 95% face images were randomly extracted from DAORL database as training set, and the remaining 5% face images as test set. The resolution of the training set and the test set is preprocessed to 32x32, and the normalized values are between 0 and 1. The initial learning rate of the network is 0.2. The random gradient descent method is used for training, with 120 pictures for each training. During the training, the exponential decay learning rate, sliding average value and regularized weight method are used to optimize the model. The default first layer convolutional kernel (c1) is 5. the same reason, c2 is 10 and c3 is 15. The DAORL database image example is shown in Fig.2.

Experimental Verification And Analysis

A. *Compared LeNet with DACNN*

The all experimental platform of this paper is in ubuntu16.04 system, processor lntel(R) Core i7-7700 CPU @3.60GHz x 8 processor and graphics card of NVIDIA GeForce GTX1080. Experiment with the ORL face database on the LeNet structure, the experimental method refers to Network Training. Changing number of c1 and c2 convolution kernels to explore the impact of the network on the recognition accuracy, as shown in TABLE I, TABLEII.

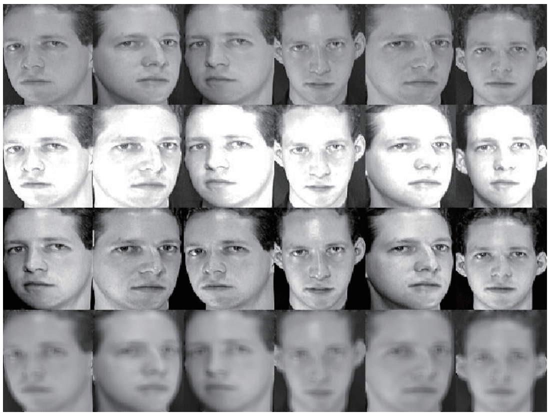

Fig. 2 The DAORL database image sample

TABLE I.
THE INFLUENCE OF C1 NUMBER ON NETWORK RECOGNITION ACCURACY

Number of c1	5	7	9	11	13
Test accuracy of LeNet(%)	92.5	97.5	95	95	95
Test accuracy of DACNN(%)	100	100	100	100	100

TABLE II.
THE INFLUENCE OF C1 NUMBER ON NETWORK RECOGNITION ACCURACY

Number of c1	12	14	16	18	20
Test accuracy of LeNet(%)	97.5	85	80	95	95
Test accuracy of DACNN(%)	100	100	100	100	100

From TABLE Ⅰ and TABLE Ⅱ,we can see that if the convolution kernel is too small, the feature extraction is insufficient during training. If the convolution kernel is overtrained, the parameters that need to be trained will increase, and a larger sample library will be needed. DACNN uses data enhancement method to enlarge the sample and train the neural network more fully under the limited samples.

B. *Influence Of Brightness And Blur On Network Recognition Accuracy*

To further illustrate the effectiveness of the DACNN method, the effects of different brightness and blur on network performance are explored. Using GaussianBlur to

Change the blurriness of the image. which is proportional to the blur radius. Using the ImageEnhance module to change the brightness of the image, which is proportional to the factor. The experimental results are shown in TABLE III.

TABLE III.
INFLUENCE OF DIFFERENT BRIGHTNESS AND BLUR ON NETWORK RECOGNITION ACCURACY

（radius, factor）	(1,1.5)	(1.5,2)	(2,2.5)	(2.5,3)	(3,3.5)
Test accuracy of DACNN(%)	100	100	100	100	98.75

As the brightness and blur of the image increasing, the network recognition accuracy will decrease. During training, the network cannot effectively extract the features of the picture for highlighted and very blurred pictures.

C. *Experimental Result*

During the training process of DACNN, the change of the loss function has a large gradient within 1000 steps. With the training, the weight becomes more and more stable and the gradient difference becomes smaller and smaller. As shown in Fig. 3. the loss function tends to 0.02 when it exceeds 3000 steps, so that a higher recognition rate can be achieved when the training exceeds 3000.

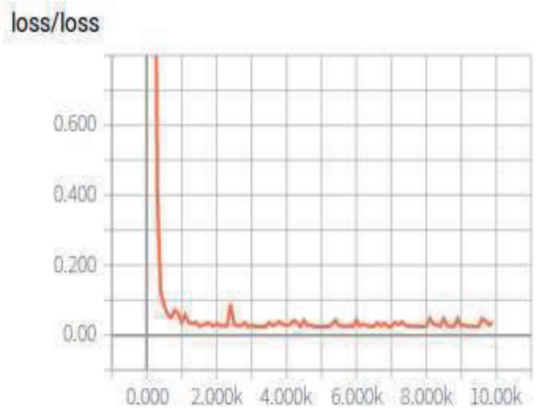

Fig.3 The variation of loss function during DACNN training

D. *Comparative Analysis*s

After the experiment, to illustrate the validity of DACNN structure, I compare it with other algorithms applied to ORL database. The correct recognition rate of Eigenface[15] algorithm is 97.5%, Fisherface[15] algorithm is 98.50%, ICA[15] algorithm is 93.75%, 2DPCA[16] algorithm is 98.30, and the DACNN algorithm is 100%, which further proves the validity of this algorithm.

Conclusions

In this paper, a convolutional neural network face recognition based on data augmentation is adopted, and the recognition accuracy in the test set generated by DAORL can be up to 100%, which has good robustness to change and blur pictures. Enhancing the database not only improves the recognition accuracy of the network, but also improves the stability of the network. As the weights of the network are initialize with a certain randomness during the training, which leads to different weights of network. Under the same parameters, the results of different trainings are different, Therefore, how to automatically find the optimal solution in training process of network is the key content of the next step. With the expansion of the database scale and the deeper network structure, it is important to automatically find the optimal solution in the training process.

References

[1] J. Murphy, "Artificial Intelligence, Rationality, and the World Wide Web," in *IEEE Intelligent Systems*, vol. 33, no. 1, pp. 98-103, Jan./Feb. 2018.

[2] T. Levi, T. Nanami, A. Tange, K. Aihara and T. Kohno, "Development and Applications of Biomimetic Neuronal Networks Toward BrainMorphic Artificial Intelligence," in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 5, pp. 577-581, May 2018.

[3] C. Ding, J. Choi, D. Tao and L. S. Davis, "Multi-Directional Multi-Level Dual-Cross Patterns for Robust Face Recognition," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 3, pp. 518-531, 1 March 2016.

[4] H. Nguyen and A. Caplier, "Local Patterns of Gradients for Face Recognition," in *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 8, pp. 1739-1751, Aug. 2015.

[5] BLEDSOE B W W.Man-machine facial recognition, panoramic research Inc[EB/OL].http://citeseerx.ist.psu.edu/showciting?cid=605724, 2010.

[6] R. Brunelli and T. Poggio, "Face recognition: features versus templates," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1042-1052, Oct. 1993.

[7] M. Turk and A. Pentland, "Eigenfaces for Recognition," in *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, Jan. 1991.

[8] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no.

7, pp. 711-720, July 1997.

[9] M. K. Fleming and G. W. Cottrell, "Categorization of faces using unsupervised feature extraction," *1990 IJCNN International Joint Conference on Neural Networks*, San Diego, CA, USA, 1990, pp. 65-70 vol.2.

[10] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry and Y. Ma, "Robust Face Recognition via Sparse Representation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227, Feb. 2009.

[11] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi and Y. Ma, "Toward a Practical Face Recognition System: Robust Alignment and Illumination by Sparse Representation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 2, pp. 372-386, Feb. 2012.

[12] E. Shelhamer, J. Long and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640-651, 1 April 2017.

[13] F. Schroff, D. Kalenichenko and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 2015, pp. 815-823.

[14] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj and L. Song, "SphereFace: Deep Hypersphere Embedding for Face Recognition," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017, pp. 6738-6746.

[15] M. Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face recognition using kernel methods," *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*, Washington, DC, USA, 2002, pp. 215-220.

[16] Jian Yang, D. Zhang, A. F. Frangi and Jing-yu Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131-137, Jan. 2004.