Human face recognition based on convolutional neural network and augmented dataset

Peng Lu, Baoye Song & Lin Xu

Published online: 27 Oct 2020.

Submit your article to this journal ☑

Article views: 4492

View related articles ☑

View Crossmark data ☑

Citing articles: 2 View citing articles ☑

Taylor & Francis
Taylor & Francis Group

Human face recognition based on convolutional neural network and augmented dataset

Peng Lu[a], Baoye Song[a] and Lin Xu[b]

[a]College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao, People's Republic of China;
[b]College of Energy and Mining Engineering, Shandong University of Science and Technology, Qingdao, People's Republic of China

ABSTRACT

To deal with the issue of human face recognition on small original dataset, a new approach combining convolutional neural network (CNN) with augmented dataset is developed in this paper. The original small dataset is augmented to be a large dataset via several transformations of the face images. Based on the augmented face image dataset, the feature of the faces can be effectively extracted and higher face recognition accuracy can be achieved by using the ingenious CNN. The effectiveness and superiority of the proposed approach can be verified by several experiments and comparisons with some frequently used face recognition methods.

1. Introduction

With the rapid development of computer science and technology, artificial intelligence (AI) is increasingly prevalent in our daily life. Among others, deep neural network (DNN) has been successfully utilized in a variety of tasks (Hinton et al., 2012; Jurgen, 2015; Liu et al., 2016; Ptucha et al., 2019; Zeng et al., 2018), e.g. character recognition, image recognition, and speech recognition, etc. In particular, face recognition based on DNN has been under intensive investigations in the past few years.

For example, a novel coupled mapping approach was proposed for the recognition of low resolution face images based on convolutional neural network (CNN) (Zangeneh et al., 2020), which is one of the architectures of DNN (Liu et al., 2016). Many face recognition experiments were carried out on the common face databases in Zhang et al. (2015), and the results showed that DNN can effectively extract the facial features based on diverse image preprocessing approaches. In Guo et al. (2015), an algorithm combining principal component analysis (PCA) with support vector machine (SVM) was presented to deal with the issue of face recognition, where PCA can not only reduce the computation load, but also promote the recognition accuracy. In Zhao et al. (2020), CNN was used to extract the facial features, then PCA was employed to reduce the dimension of the obtained features. Moreover, joint Bayesian algorithm was applied for the face recognition. The final experiment results can validate the superiority of the hybrid approach on the

recognition accuracy. Nevertheless, it is worth noting that most of DNN-based face recognition methods are usually developed based on a large original dataset (Guo & Zhang, 2019; Sun & Meng, 2016; Zhang et al., 2015). Obviously, a large original dataset can provide more features of the face images, but it is usually difficult to be obtained in comparison with a small original dataset. As a result, some of the existing successful methods could lead to poor performance on the small original dataset. Besides, though a larger original dataset can bring higher accuracy of the model and stronger generalization ability of the network, the data labelling of a large original dataset is really a tedious and time-consuming work. Therefore, from a practical point of view, it is a promising topic to develop the DNN-based face recognition methods on the small original dataset (Huang & Mu, 2014).

Motivated by above-mentioned observations and considerations, a new approach combining convolutional neural network with augmented dataset is developed for human face recognition in this paper. The main contributions of this paper can be summarized as threefold. (1) The small original dataset is augmented to be a large dataset by using several transformations of the face images. (2) Based on the augmented human face dataset, the face recognition is implemented via an ingenious CNN, which is robust to the image transformations. (3) Several experiments are carried out on a common face database, and the superiority of the new approach can be confirmed by comparing with some of the frequently used methods.
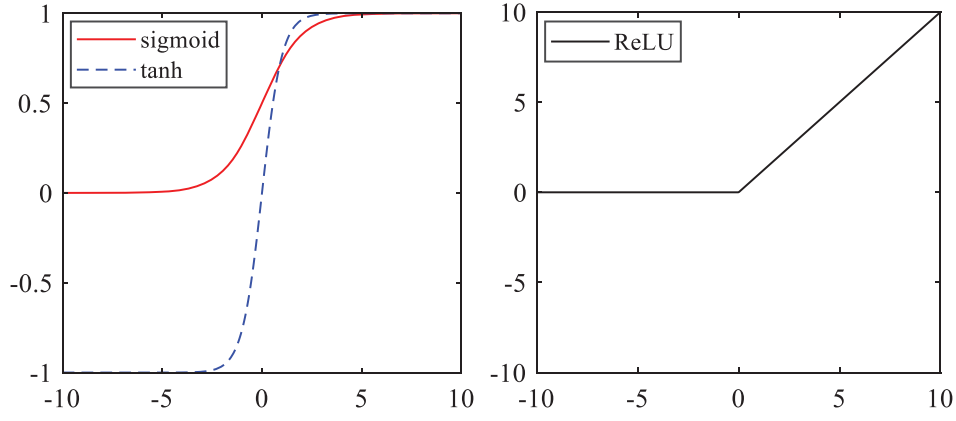

CONTACT Baoye Song ✉ songbaoye@sdust.edu.cn

Figure 1. Activation functions.

The remainder of the paper can be outlined as follows. A preliminary of CNN is briefly introduced in Section 2 for the sake of completeness. Section 3 is devoted to the face recognition method, where the model of CNN is sketched and the augmentation method of face images is described in this section. In Section 4, some experiments are carried out to verify the effectiveness of the augmented dataset and the superiority of the proposed approach. Finally, the conclusion is drawn in Section 5.

2. Preliminary of convolutional neural network

Convolutional neural network (CNN) is a kind of neural network with convolutional layers (LeCun et al., 1989, 1998). In general, CNN contains two kinds of hidden layers, i.e. convolutional layers and pooling layers, which are usually arranged alternately in the neural network (Wang et al., 2012; Zeiler & Fergus, 2014).

Similar to biological neural network, the connection weights of CNN can be shared in the whole neural network, which can not only reduce the amount of the connection weights, but also simplify the complexity of the network model. Thus, the training time of CNN can be remarkably shortened in most cases. In particular, when an image is the input of CNN, the image can be put into the neural network directly to avoid several complicated works, such as feature extraction and data reconstruction. Owing to the advantages of weight sharing, pooling and local receptive field, CNN has a robust performance on several image transformation operations, e.g. translation, rotation, and scaling. For the sake of completeness, the preliminary of CNN is briefly introduced in the rest of this section.

2.1. Activation function

The performance of a neural network is closely related to not only its structure but also the adopted activation function, which is usually selected as a nonlinear function to deal with some complex issues. Three frequently used activation functions in CNN are **sigmoid**, *hyperbolic tangent* (*tanh*) and *rectified linear unit* (*ReLU*) (Bengio et al., 2016; Lin & Shen, 2018), which can be formulated as follows and illustrated in Figure 1.

$$\text{sigmoid}: f(x) = \frac{1}{1 + e^{-x}}, \tag{1}$$

$$\text{tanh}: f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \tag{2}$$

$$\text{ReLU}: f(x) = \max(0, x). \tag{3}$$

2.2. Back propagation algorithm

Back propagation (BP) (LeCun et al., 1989) algorithm is one of the most frequently used algorithms to train a neural network, and the mapping of the input and output data is actually a nonlinear optimization problem of the connection weights. Based on the gradient descent (GD) of BP algorithm (LeCun et al., 1998), the connection weights of neural network can be updated iteratively by minimizing the mean square error (MSE) between the real and expected values of the output. Here, the MSE, which is usually defined as the cost function in the training of neural network, can be expressed as:

$$E(\mathbf{W}, \mathbf{B}) = \frac{1}{N_L} \sum_{i=1}^{N_L} (a_{iL} - t_{iL})^2, \tag{4}$$

where **W** and **B** denote, respectively, the weight and bias matrices to be optimized in the neural network; a_{iL} and t_{iL} indicate, respectively, the real and expected output values of the ith neuron in the output layer with N_L neurons.

In the neural network shown in Figure 2, the output of the ith neuron in the lth layer can be calculated as follows:

$$a_{il} = f_{il} \left(\sum_{j=1}^{N_{l-1}} w_{ijl} a_{j,l-1} + b_{il} \right), \tag{5}$$
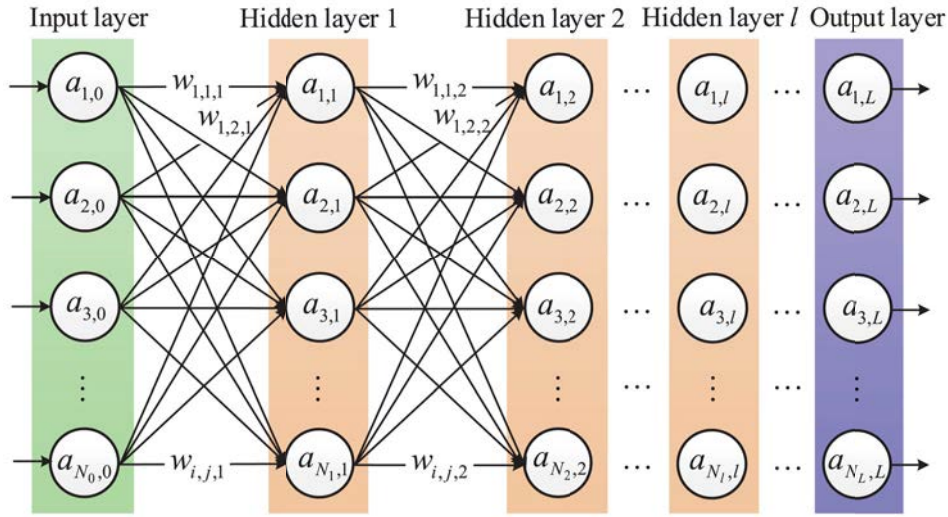
Figure 2. Structure of neural network.

where $f_{il}(\cdot)$ and b_{il} stand for the activation function and bias of a_{il}, respectively; w_{ijl} denotes the connection weight between the ith neuron of the lth layer and the jth neuron of previous layer; l indicates the layer number of the L-layer neural network; and N_l is the neuron amount of the lth layer. Based on the forward propagation of the input and the cost function defined above, the weights and biases of the neural network can be updated iteratively as follows:

$$w_{ijl}(k+1) = w_{ijl}(k) + \triangle w_{ijl}(k), \tag{6}$$

$$b_{il}(k+1) = b_{il}(k) + \triangle b_{il}(k), \tag{7}$$

where $\triangle w_{ijl}$ and $\triangle b_{il}$ can be calculated according to the chain rule of calculus as follows:

$$\triangle w_{ijl} = -\eta \frac{\partial E}{\partial w_{ijl}} = -\eta \frac{\partial E}{\partial a_{il}} \frac{\partial a_{il}}{\partial w_{ijl}}, \tag{8}$$

$$\triangle b_{il} = -\eta \frac{\partial E}{\partial b_{il}} = -\eta \frac{\partial E}{\partial a_{il}} \frac{\partial a_{il}}{\partial b_{il}}, \tag{9}$$

where η is the learning rate that can control the rate of gradient descent. Note that the above-mentioned chains are not expanded in detail for the sake of simplicity. The interested readers can refer to Ranganathan et al. (2019) and Hagan et al. (2002) for more details on BP algorithm.

2.3. Convolution operation

Convolution is a kind of mathematical operation that has been widely used in image processing. The result of convolution can be sorted as three modes (Bengio et al., 2016), i.e. the modes of Full, Same and Valid, which can be utilized in different occasions. For example, Valid mode is usually applied for forward propagation to facilitate the feature extraction of image, and Full mode is

often employed in the back propagation to obtain the optimal weights.

In the convolution operation, the operation of edge zeroing is implemented for the input image, where the layer amount of the edge can be determined according to the size of the convolution kernel (Lawrence et al., 1997). The purpose of edge zeroing is to ensure the rationality of the results, i.e. the elements of the input image and the convolution kernel can be weighted and summated sequentially. Additionally, the convolution kernel should be turned around and flipped up and down as shown in Figure 3, where the kernel is actually rotated 180 degrees around the centre.

It is worth noting that convolution operation can achieve sparse multiplication and parameter sharing, which can compress the dimension of the input data. In comparison with DNN, it is not necessary for CNN to provide connection weights separately for all neurons of the input data. Actually, CNN can be regarded as a common feature extraction process like most neural networks used for feature extraction.

2.4. Receptive field

In CNN, the receptive field is the local connection field of a neuron in the hidden layer (Alonso & Chen, 2009). Suppose that the input of the neural network is an image with 100×100 pixels and there are 100 neurons in the hidden layer, there will be $100 \times 100 \times 100$ connection weights between the input and the hidden layers if every pixels of the image are connected to all neurons of the hidden layer as shown in Figure 4(a). There is no doubt that the huge computation load will decrease the training efficiency of the neural network. In contrast, if each hidden neuron is connected to a local field of the input image
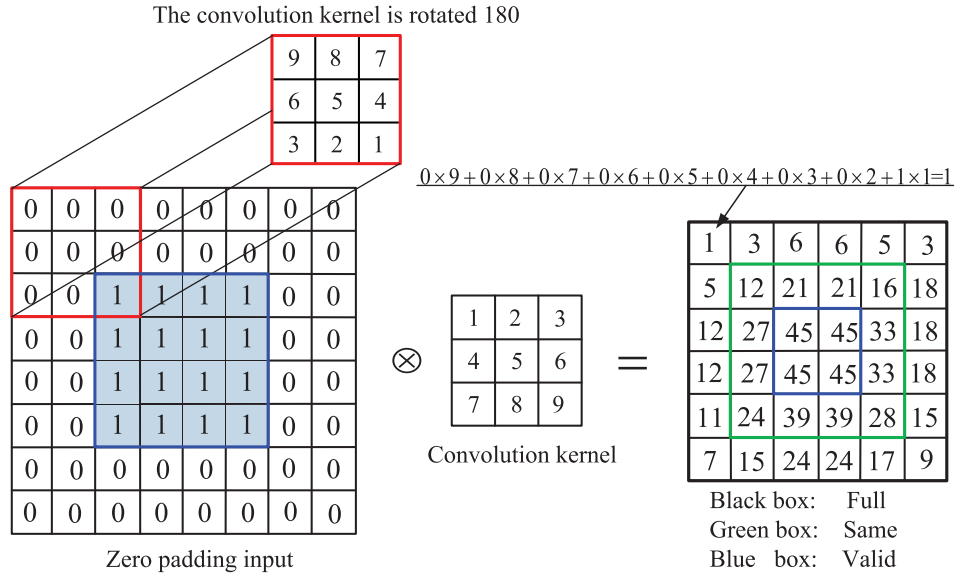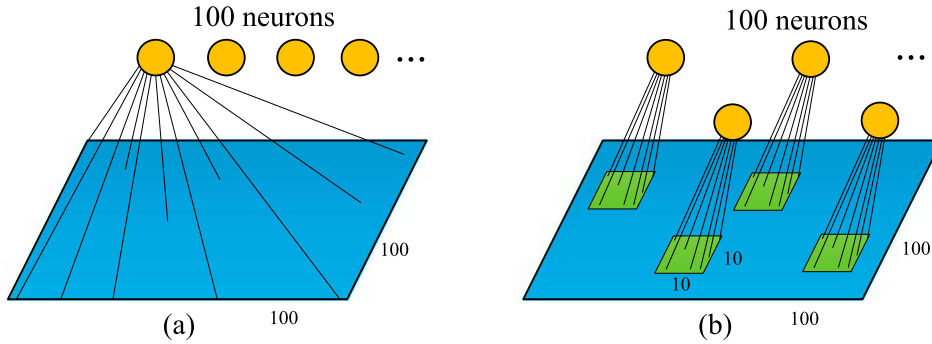
The convolution kernel is rotated 180

$0 \times 9 + 0 \times 8 + 0 \times 7 + 0 \times 6 + 0 \times 5 + 0 \times 4 + 0 \times 3 + 0 \times 2 + 1 \times 1 = 1$

Convolution kernel

Black box: Full
Green box: Same
Blue box: Valid

Zero padding input

Figure 3. Convolution operation.

100 neurons

100 neurons

(a)

(b)

Figure 4. Receptive field.

(e.g. 10 × 10 pixels) as shown in Figure 4(b), the amount of connection weights will be reduced to 10 × 10 × 100, which is 1/100 of the full connection case.

In practice, the connection weights can be further reduced by using the shared weight method, i.e. all neurons have the same weights in one convolution kernel. Thus, the amount of connection weights can be reduced from 10 × 10 × 100 to 10 × 10. Accordingly, the training speed of the neural network can also be significantly promoted.

2.5. Pooling

The pooling layers, which are usually located behind the convolutional layers, are mainly used to compress the output feature data of the convolutional layers. After the pooling layer, the improved output results can reduce the likelihood of over-fitting in the neural network. Besides, the feature of image can be further extracted through pooling operation without influencing on the information acquisition of the image.

Actually, pooling is a reduction processing of the image, which can be classified as mean-pooling, max-pooling, overlapping-pooling, stochastic-pooling, and global average pooling (He et al., 2015; Krizhevsky et al., 2017). For instance, mean-pooling can extract the average value of the feature points and has the effect of maintaining the relative background; while max-pooling can extract the maximum value of the feature points and achieve better texture extraction. Specifically, for the mean-pooling, if a feature map with size of 4 × 4 is sampled by using a kernel with size of 2 × 2, the output is a feature map with size of 2 × 2 as shown in Figure 5.

3. Face recognition method

3.1. CNN model for face recognition

In this paper, a CNN model is developed to improve the accuracy of face image classification. The structure of the model is similar to the classical LeNet-5 model, but they are different on some parameters of the model, such as input data, network width and full connection layer. The
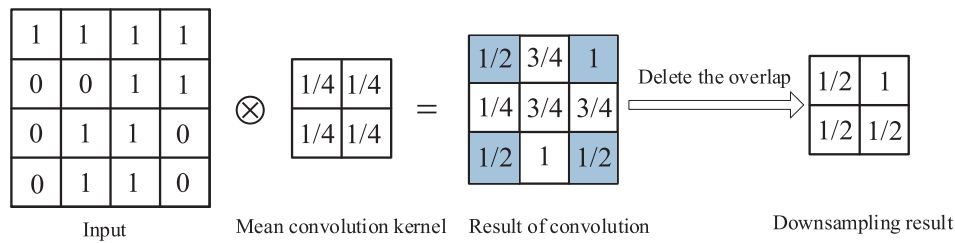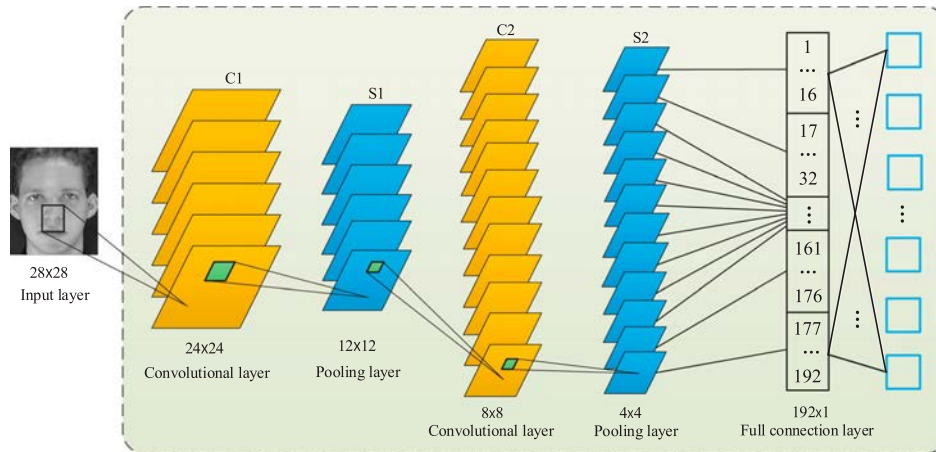
Figure 5. Pooling operation.

Figure 6. CNN model for face recognition.

developed CNN is composed of two convolutional layers (C1 and C2) and two pooling layers (S1 and S2). These layers are arranged alternately in the form of C1-S1-C2-S2 as sketched in Figure 6.

There is only one feature map in the input layer, which is used to put the normalized face image into the CNN model. C1 is the first convolutional layer that includes 6 feature maps, in which each neuron is convoluted with a randomly generated convolution kernel with size of 5×5. S1 is the first pooling layer, whose output is 6 feature maps calculated based on the output of previous layer. Each element in the feature map is connected with the mean convolution kernel of the corresponding feature map in C1 layer, and the receptive fields of the elements will not overlapped with each other. C2 and S2 are, respectively, the second convolutional layer and pooling layer, both of which have 12 feature maps and similar calculation steps with their previous counterparts. Moreover, a fully connected single-layer perceptron is placed between the S2 layer and the output layer. As shown in Figure 6, the final output is a 40-dimensional vector for the face recognition of 40 individuals, where the *sigmoid* function is used for the multi-label classification.

3.2. Dataset and its augmentation

Olivetti Research Laboratory (ORL) face dataset is a collection of 400 human face images from 40 individuals (ATT Laboratories Cambridge, 2005). Each individual devotes 10 face images in diverse states to the dataset. The size of each image is 92×112 pixels, and each of the images is saved as the file types of BMP and PGM. ORL face dataset is a widely used standard face dataset that is more easier to be labelled than others such as MIT or Yale face datasets. However, the amount of images is not abundant to train the deep neural network for accurate face recognition. To deal with this problem, the image amount of the dataset is augmented by using four data augmentation methods, including horizontal flip, shift, scaling, and rotation as shown in Figure 7.

Actually, the dataset can be augmented tremendously by tuning the parameters of the augmentation methods, see Figure 8 for example. In this paper, the dataset is augmented by 1000 times after aforementioned operations. Then, the images are scaled, normalized and labelled before they are put into the face recognition system. It can be predictable that the augmented dataset can not only reduce the probability of over-fitting but also improve the robustness of the system.

Remark 3.1: It is obvious that a large original dataset can provide abundant image features for the human face recognition problem. Unfortunately, however, such an ideal database is difficult to obtain in practical applications. By using the transformation of the face images, the small image dataset can be tremendously augmented
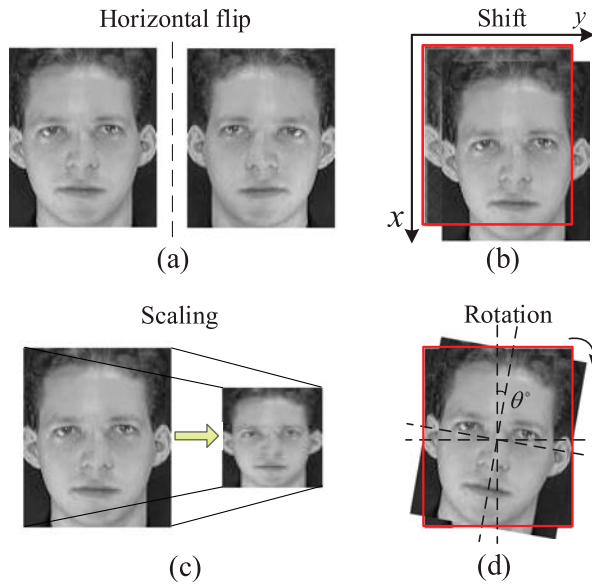
Horizontal flip

Shift

(a)   (b)

Scaling

Rotation

(c)   (d)

Figure 7. Four methods for data augmentation.

to be a large one. As such, more image features can be extracted to train the classifier so as to achieve better face recognition result.

4. Experiment results and analyses

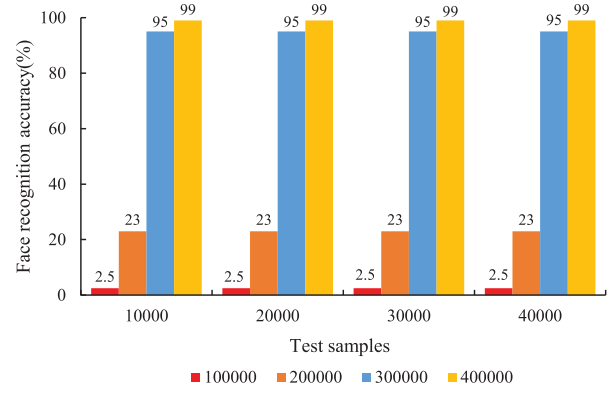In this section, some experiments are implemented to evaluate the performance of the developed face

Figure 9. Face recognition accuracy with different amounts of training and test samples.

recognition method, and the superiority of the method can be verified in comparing with some frequently used methods. The experiments are carried out using MATLAB 2018a on a PC with Intel Core i7-6700 CPU.

To test the effect of the augmented dataset, the face recognition network is trained using different amounts of samples, i.e. 100,000, 200,000, 300,000, and 400,000, where all batch sizes and epochs are set as 50 and 1, respectively. Meanwhile, different amounts of test samples are adopted in the test of the neural network, see Figure 9 for more details. It can be observed that the face
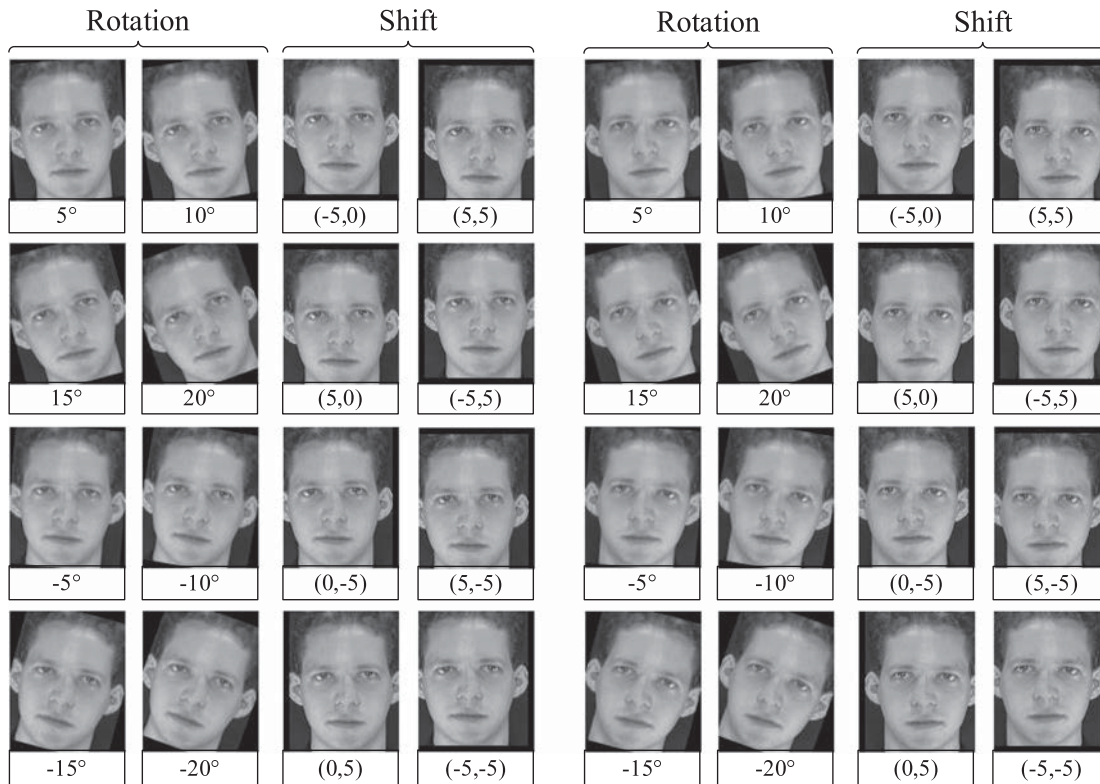

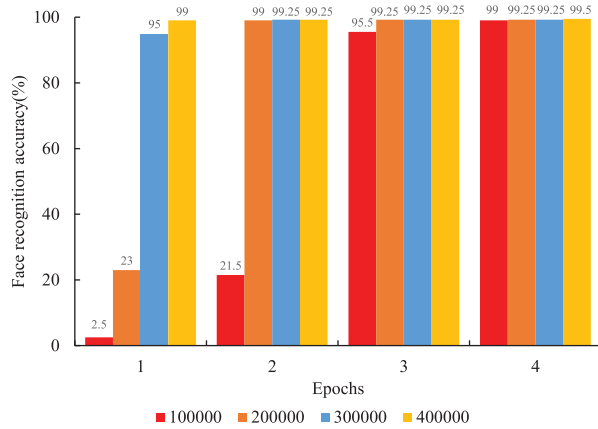**Figure 8.** Data augmentation for the first face.

Figure 10. Face recognition accuracy with different amounts of training samples and epochs.

recognition accuracy can be promoted with the increasing of training samples, simultaneously maintains stable for different amounts of test samples. This means that the augmented dataset can introduce abundant sample features, which can enhance the training of network and result in high face recognition accuracy.

As mentioned above, the face recognition network is robust to different amounts of test samples, so that the test sample can be kept constant in the following experiments. Now, the network is trained using different amounts of samples and epochs, and the results are summarized in Figure 10. It is clear that the face recognition

accuracy will increase with more training samples in the cases of different epochs. Meanwhile, more epochs will also promote the accuracy in the case of different training samples. That is the face recognition accuracy is positively correlated with the amounts of epochs and training samples. Actually, this conclusion can also be confirmed in the experiments using more epochs and training samples.

MSE is usually used for the cost function of neural network. In the back propagation of error, MSE is the expectation of the squared error between the real and expected values of the output. The small MSE means high accuracy of the network model. To analyse the MSE of the face recognition network, some experiments are implemented using different amounts of training samples (i.e. 100,000, 200,000, 300,000, and 400,000) and epochs (i.e. 1, 2, 3, and 4) with the batch size of 50. As seen in Figure 11, the MSE will decrease with the increasing of training iterations in all cases of training samples. Additionally, the results of MSE can further verify the conclusions of the previous experiments.

To further verify the superiority of the proposed approach, the results of the approach are compared with some other face recognition methods in the literature based on ORL face dataset, including ANN (artificial neural network), PCA+ANN, PCA+SVM, Wavelet+SVM, and Wavelet+SVM (Gumus et al., 2010; Guo et al., 2015), and their face recognition accuracies are listed in Table 1. It is clear that the SVM-based methods can outperform the
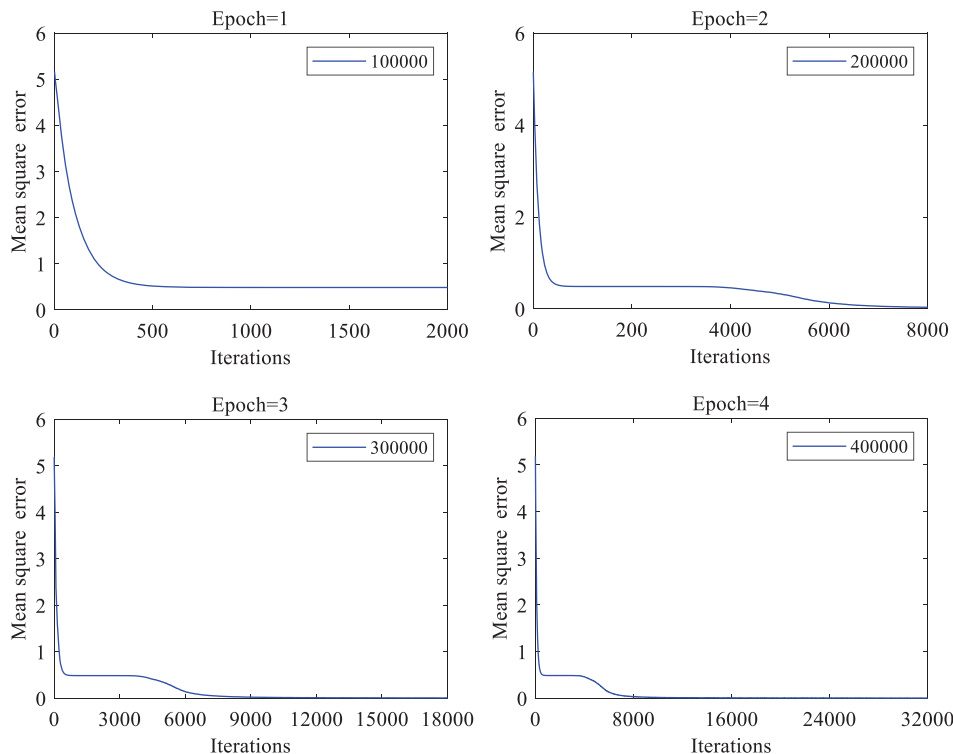

Figure 11. MSE with different amounts of training iterations and samples.

Table 1. Comparison of face recognition accuracy.

Method	Accuracy (%)
CNN	99.5
ANN	80.3
PCA+ANN	91.0
PCA+SVM	97.4
Wavelet+SVM	98.1
Wavelet+PCA+SVM	98.0

ANN-based methods on the face recognition accuracy of the ORL face dataset. While, the proposed approach, which combines CNN with the augmented face dataset, can achieve a higher face recognition accuracy than others. As such, it can be concluded that the augmented face dataset can promote the performance of face recognition owing to the obtainable abundant features.

5. Conclusion

In this paper, a new approach has been developed to deal with the problem of human face recognition on small original dataset. The original small dataset is augmented to be a large dataset by using the transformations of the face images, such as flip, shift, scaling, and rotation. Based on the remarkably augmented face dataset, the face recognition can be effectively implemented via an ingenious CNN. Several experiments are carried out to verify the effectiveness of the augmented dataset, and the superiority of the new approach can also be confirmed in comparison with some of the frequently used face recognition methods.

Actually, the proposed strategy is an economic method to augment the dataset, and can be applied to a variety of fields related to data-based training and learning. Our future work will focus on the application of the data augmentation approach on some more complex problems, e.g. signal processing, image recognition, and image-based fault detection (Chen et al., 2018; Huang et al., 2020; Song et al., 2019; Wang & Yang, 2019; Xu et al., 2018; Zeng et al., 2017), etc.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

References

Alonso, M., & Chen, Y. (2009). Receptive field. *Scholarpedia*, *4*(1), 5393. https://doi.org/10.4249/scholarpedia.5393

ATT Laboratories Cambridge (2005). *The ORL database of faces*. http://www.cam-orl.co.uk/facedatabase.html

Bengio, Y., Goodellow, I., & Courville, A. (2016). *Deep learning*. MIT Press.

Chen, J., Zhou, D., Wang, Y., Fu, H., & Wang, M. (2018). Image feature extraction based on HOG and its application to fault diagnosis for rotating machinery. *Journal of Intelligent & Fuzzy Systems*, *34*(6), 3403–3412. https://doi.org/10.3233/JIFS-169521

Gumus, E., Kilic, N., Sertbas, A., & Ucan, O. N. (2010). Evaluation of face recognition techniques using PCA, wavelets and SVM. *Expert Systems with Applications*, *37*(9), 6404–6408. https://doi.org/10.1016/j.eswa.2010.02.079

Guo, L., Leng, J., Mei, W., Kong, X., Liao, Y., & Liao, H. (2015). Research on human face recognition technology based on PCA and SVM. *Journal of Hubei University for Nationalities (Natural Science Edition)*, *33*(2), 193–196+214. https://doi.org/10.13501/j.cnki.42-1569/n.2015.06.020

Guo, G., & Zhang, N. (2019). A survey on deep learning based face recognition. *Computer Vision and Image Understanding*, *189*, 102805. https://doi.org/10.1016/j.cviu.2019.102805

Hagan, M. T., Demuth, H. B., Beale, M. H., & Jesús, O. D. (2002). *Neural network design*. China Machine Press.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *37*(9), 1904–1916. https://doi.org/10.1109/TPAMI.2015.2389824

Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A. R., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Nguyen, T., Sainath, T. N., & Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, *29*(6), 82–97. https://doi.org/10.1109/MSP.2012.2205597

Huang, X., Liu, Z., Zhang, X., Kang, J., Zhang, M., & Guo, Y. (2020). Surface damage detection for steel wire ropes using deep learning and computer vision techniques. *Measurement*, *161*, 107843. https://doi.org/10.1016/j.measurement.2020.107843

Huang, H., & Mu, Z. (2014). Review face recognition on small sample size problem. *Software*, *35*(3), 167–169. https://doi.org/10.3969/j.issn.1003-6970.2014.03.056

Jurgen, S. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Krizhevsky, A., Sutskever, I., & Hinton, G. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

Lawrence, S., Giles, C. L., & Tsoi, A. C. (1997). Face recognition: A convolutional neural-network approach. *IEEE Transactions on Neural Networks*, *8*(1), 98–113. https://doi.org/10.1109/72.554195

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, *1*(4), 541–551. https://doi.org/10.1162/neco.1989.1.4.541

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Lin, G., & Shen, W. (2018). Research on convolutional neural network based on improved Relu piecewise activation function. *Procedia Computer Science*, *131*, 977–984. https://doi.org/10.1016/j.procs.2018.04.239

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2016). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*, 11–26. https://doi.org/10.1016/j.neucom.2016.12.038

Ptucha, R., Petroski Such, F., Pilla, S., Brockler, F., Singh, V., & Hutkowski, P. (2019). Intelligent character recognition using fully convolutional neural networks. *Pattern Recognition*, *88*, 604–613. https://doi.org/10.1016/j.patcog.2018.12.017

Ranganathan, S., Gribskov, M., Nakai, K., & Schönbach, C. (2019). Delta rule and backpropagation. *Encyclopedia of Bioinformatics and Computational Biology*, *1*, 621–633. https://doi.org/10.1007/0-387-22733-4_9

Song, B., Xu, J., & Xu, L. (2019). Multilevel inverter fault diagnosis based on WPT-PCA-NN algorithm. *Journal of Shandong University of Science and Technology (Natural Science)*, *38*(1), 111–120. https://doi.org/10.16452/j.cnki.sdkjzk.2019.01.013

Sun, J., & Meng, F. (2016). Face recognition based on deep neural network and weighted fusion of face features. *Journal of Computer Applications*, *36*(2), 437–443. https://doi.org/10.1016/j.compind.2018.12.001

Wang, T., Wu, D. J., Coates, A., & Ng, A. Y. (2012). End-to-end text recognition with convolutional neural networks. In *Proceedings of 2012 21st international conference on pattern recognition (ICPR 2012)*, Tsukuba, Japan, November 2012 (pp. 3304–3308).

Wang, Z., & Yang, R. (2019). Fault diagnosis of rotating machinery gearbox based on random forest algorithm. *Journal of Shandong University of Science and Technology (Natural Science)*, *38*(5), 104–112. https://doi.org/10.16452/j.cnki.sdkjzk.2019.05.013

Xu, L., Cao, M., Song, B., Zhang, J., Liu, Y., & Alsaadi, F. E. (2018). Open-circuit fault diagnosis of power rectifier using sparse autoencoder based deep neural network. *Neurocomputing*, *311*, 1–10. https://doi.org/10.1016/j.neucom.2018.05.040

Zangeneh, E., Rahmati, M., & Mohsenzadeh, Y. (2020). Low resolution face recognition using a two-branch deep convolutional neural network architecture. *Expert Systems with Applications*, *139*, 112854. https://doi.org/10.1016/j.eswa.2019.112854

Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convoluional network. *Lecture Notes in Computer Science*, *8689*(1), 818–833. https://doi.org/10.1007/978-3-319-10590-1

Zeng, N., Zhang, H., Li, Y., Liang, J., & Dobaie, A. M. (2017). Denoising and deblurring gold immunochromatographic strip images via gradient projection algorithms. *Neurocomputing*, *247*, 165–172. https://doi.org/10.1016/j.neucom.2017.03.056

Zeng, N., Zhang, H., Song, B., Liu, W., Li, Y., & Dobaie, A. M. (2018). Facial expression recognition via learning deep sparse autoencoders. *Neurocomputing*, *273*, 643–649. https://doi.org/10.1016/j.neucom.2017.08.043

Zhang, Z., Li, J., & Zhu, R. (2015, October). Deep neural network for face recognition based on sparse autoencoder. In *Proceedings of 2015 8th international congress on image and signal processing (CISP)* (pp. 594–598). Shenyang, China.

Zhao, F., Li, J., Zhang, L., Li, Z., & Na, S. G. (2020). Multi-view face recognition using deep neural networks. *Future Generation Computer Systems*, *111*, 375–380. https://doi.org/10.1016/j.future.2020.05.002