Human Identification Based on Deep Feature and Transfer Learning

Chen Li[1] , Zhiqiang Weng[1] , Qianli Wang[1] , Ke Xiao[1] , Wei Song[1] ,Hui Zhou[1]

[1]School of Computer Science, North China University of Technology

No.5 JinyuanzhuangRoad, Shijingshan District,

Beijing 100144, China

Corresponding: lichen@ncut.edu.cn

Abstract

Biometric based identity authentication has attracted much attention due to its unique advantages. Among all the biometric which can be used for authentication, human face based methods have been the most popular research area in both identity authentication and recognition. However, traditional method may result in poor performance when conducting face recognition under uncontrolled environmental. Deep network provides a more proper way to extract distinctive features for face recognition, however the performance of most deep network is usually limited by the number of training samples. Accordingly, this paper proposes a deep convolutional neural network combining with the idea of transfer learning and sparse representation to combat the disadvantage of traditional CNN on small sample task while simplifying the computational complexity. Abundant experimental results in different database show that compared with traditional method, our proposed method achieves higher and promising recognition rate.

CCS concepts

Security and privacy → Human and societal aspects of security and privacy

Keywords

Face recognition; Deep feature; Convolution Neural Networks (CNN); Transfer learning.

1. INTRODUCTION

Traditional identification methods usually use passwords, certificates or certain specific knowledge to set the user's access to the information system. However, such traditional identification methods have many shortcomings: easy to lose or forget, easy to be cracked [1]. Therefore, identity authentication with biometrics as a key technology has been vigorously developed. Biometric technology has become one of the safest ways for identification because of its stability and unforgeability. Among all the biometric features, the acquisition of face is not only convince but also non-intrusive, which makes face recognition a good candidate for both research and commercial application [2].

As an important biometric technology, face recognition has broad application prospects in information security, access control, monitoring, smart card, law enforcement, human-computer interaction, entertainment [3]. Therefore, face recognition in lab condition has attracted many researchers and been studied for years [4]. Many representative algorithms has been proposed including: Principal Components Analysis (PCA), Local Binary Pattern (LBP), and Support Vectors Machine (SVM) and so on. However, in practical applications, face recognition technology is still not broadly used [5]. The main reason is that when leaving the lab condition, many uncontrolled factors may occur, including illumination variation, occlusion, self occlusion, expression, age, small samples and etc. And the traditional methods shows not very excellent performance when combating these factors. With the development of machine learning, face recognition technology has also made a breakthrough development when conducting a great number of training samples [6]. However, for the scenario when only small training samples can be obtained for each subject, such as: legal implementation, passport verification and etc.[7], only these limited numbers of images can be used to train the deep network [8].

To combat this problem, in this paper, a face identification method based on the combination of deep feature and transfer learning is presented. To improve the training efficiency of the deep network, the sparse presentation concept is also introduced. Besides, database augmentation are conducted for pre-training. On the basis of pre-training, the obtained deep network can be applied to the new but small sample data by transfer learning,

The structure of the following paper is shown as follow: Section 2 provides the basic theoretical model of deep learning neural networks, and the basic framework of convolutional neural network are introduced. The classification method used in traditional algorithms and the theoretical basis of migration learning are briefly introduced. The deep convolutional neural network model designed and implemented in the paper, the training method and how to extract the feature vector of the image using the trained model are described in detail in Section 3. And in the last Section, this paper provides the face database used in the experiment and the preprocessing of the face database, and applies the network designed in this paper to face recognition task.

BDIOT 2018, October 24–26, 2018, Beijing, China.

ACM ISBN 978-1-4503-6519-2/18/10…$15.00

https://doi.org/10.1145/3289430.3289464

2. TRANSFER LEARNING BASED DEEP FEATURE EXTRACTION

Convolutional neural networks (CNN) is a special deep neural network model. The connection between its neurons is not completely connected. And between the neurons of the same layer, the connection weight is shared. Its weight-sharing network structure makes it more similar to biological neural networks, which can help reduce the complexity of the network model and make the network structure very suitable for image recognition task. In this paper, on the basic of CNN, a transfer learning based deep feature extraction framework (T-DFE) is presented for conquering the disadvantage of traditional CNN on small sample task.

2.1. Deep network training and transfer

In this paper, we trained a simply convolutional neural network to combine with the transfer learning concept for deep feature extraction. The network structure includes: an input layer, three convolution layer (each with a pooling layer), and then a fully connection layer. The specific structure and parameters are shown in Figure 1.

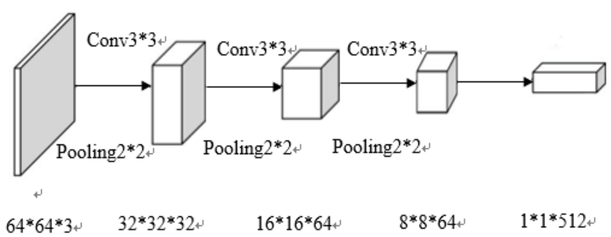

Figure 1. Sparse deep network

Considering that in many specific face recognition scenarios, acquiring abundant face images is a quite challenging even impossible task, in this paper we propose to conduct face recognition combining with transfer learning concept. The goal of Transfer Learning is to use the knowledge learned from one environment to help with learning tasks in other new environment. Considering the similarity of training and testing task, the trained deep network structure and model parameters can be transfer to the new model in a certain way to accelerate and optimize the learning efficiency of the model. The scheme is that an extended face database is constructed from the augmentation of AR, FERET, or IMM database. Then it is used to pre-train the deep convolutional neural network. While training the deep model, the dropout procedure is applied to each hidden layer to reduce over-fitting and simplify the training calculation.

Then to conduct small sample face recognition, the obtained structure and parameters are applied for the recognition of other original database with small sample by transfer learning, as shown in Figure 2.

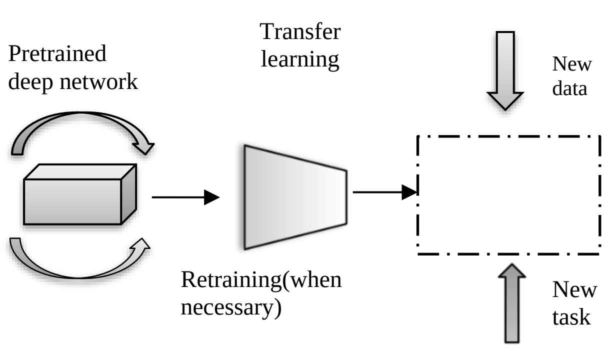

Figure 2. Transfer Learning

Besides, to verify the performance of the deep network, the chosen test database has completely no repetitive image with the training database. Although, the effect of transfer learning is not as good as full retraining, but it can help conquering the insufficient of samples as well as reduce the computational complexity, which is very suitable for solving small sample problems.

2.2. Feature extraction and Recognition

In the presented deep network, each higher layer can abstract higher level or deeper features from the lower level feature. For image classification tasks, the neural network automatically removes irrelevant features such as background, image position, etc., and then magnifies the more recognizable features such as shapes, edges, and so on. Hence, each layer in the network structure before the classifier can be regarded as a general feature extractor.

To further simplify the proposed deep network and enhance its flexibility, the softmax layer is removed when conducting transfer learning. Instead, the classification task is conducted by a linear classifier. In this paper, the Euclidean distance is applied. Euclidean distance is widely used in image similarity computing. The output of the fully connection layer can be regarded as the deep features of the input face images since that the deep network has distinguished feature extraction ability. Then the classification task can be completed by combining the deep feature with the linear classifier, so that the deep CNN model can well complete the new recognition task.3.

3. EXPERIMENT

In this paper, the ORL, IMM and AR face database are all used for recognition experiments to verify the performance of the proposed method. The details of these databases and the experiments are introduced in this section.

3.1. Face database

3.1.1 ORL face database

The ORL Face Database consists of 400 face images from 40 people, each with 10 images. The images in ORL database include facial expressions variation (laugh/no smile, blink/close eyes) . And pose variation. Figure 3 shows partial sample images of the ORL face database.

Figure 3. ORL Face Database Samples

3.1.2 IMM face database

The IMM Face Database includes 240 images from 40 different subjects. The gender distribution is 7 women and 33 men. Each subject has 6 images with pose variation, expression variation as well as different illumination condition. The image resolution is a 640-pixel ×480-pixel JPEG format. Figure 4 are some sample images of the IMM face database.

Figure 4. IMM Face Database

3.1.3 FERET face database

The FERET face database are obtained from 200 different subjects, each with 7 images. All the images are gray-scale image with different poses, different illuminations and different facial expressions. FERET face database is one of the most widely used face databases in the field of face recognition. In this paper the FERET-40 face database represent the first 40 categories of FERET. It is chosen to be the training set. Figure 5 shows the face images of two people from FERET.

Figure 5. FERET Face Database

3.1.4 AR face database

The AR database are obtained under different angles, different illuminations, different facial expressions, with or without masks (sunglasses, scarves), and the percentage of the occluded images are 46.16%. 40 subject from AR database are used in this experiment and for each person 14 images with different postures, facial expressions, and changes in illumination are remained. The total number is 560 grayscale images.

Figure 6. AR Face Database

3.2 Experimental results

At first, each face database are extended through illumination variation. Then the image number of each subject from the AR face database are augmented to 700. The image number of each subjects from FERET-40 face database are augmented to 700 images and the image number of each subjects from IMM database are augmented to 720 images. They all been used to train the network model respectively. Then these three deep network models are migrated to ORL face database for fine-tuning learning. Because of the influence of dropout, the eigenvalues of the same image extracted by neural network are slightly different each time. 5 times feature extraction are conducted for each model and the average recognition rate are calculated as the final accuracy. After comparing the accuracies, the model trained by FERET-40 are chosen as the final network model .In order to shown the effectiveness of our presented method, comparison experiments are conducted in this part.

In the first experiment, the deep convolutional neural network model pre-trained by the FERET-40 face database are chosen as the basic network. Then the basic network are transfered to conduct recognition of face images from original ORL as testing. During which, the output of the full connection layer in the network is applied as the feature vector of the image, and the recognition task is completed by the Euclidean distance based classification method. To show the effectiveness of the proposed method CNN-T, LBP is selected for comparison. The recognition result of both CNN-T and LBP as are shown with CMC curves in Fig7.

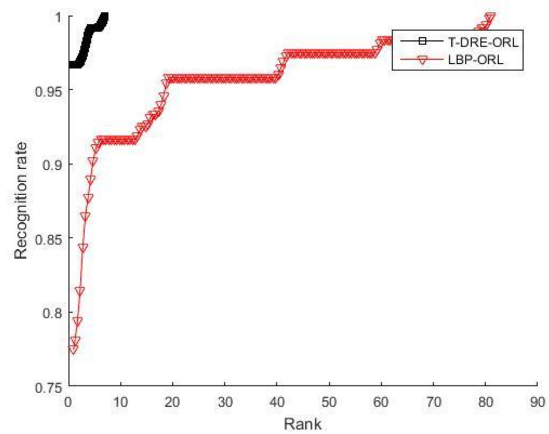

Figure 7. Recognition rate comparison of T-DRE and LBP

As shown, the recognition rate of the presented T-DRE is 97.25% and it is obviously higher than that of LBP in original ORL database.

In the second experiment, we use the same deep convolutional neural network pre-trained by the same database. Then transfer it to conduct recognition of face images from original IMM database. LBP is chosen as the comparison. The recognition results are shown in Figure 8. As shown, the rank1 recognition rate of T-DRE is 97.6667%， and the rank1 recognition rate of LBP is 81.63%. It can be seen that although the face images in IMM database are capture with expression, illumination and pose variation, the presented T-DFE obtains satisfied result.

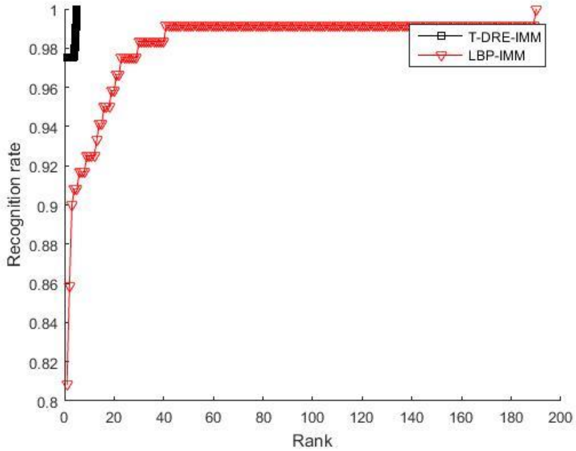

Figure 8. Recognition rate comparison of T-DRE and LBP in IMM database

In the third experiment, to remain the same experiment scheme for comparison, 40 subjects are randomly selected from the original AR database. The pre-trained deep network are transfered to conduct recognition of face images from AR-40 database. For each subject, 10images are selected as training dataset and the remained 4 images are selected as testing dataset. The recognition results are still compared with the result of LBP. As shown in Fig.8, the presented T-DFE achieves 83.75% rank-1 recognition rate, while LBP achieves 77.5% rank-1 recognition rate. Our presented method shows obvious advantages.

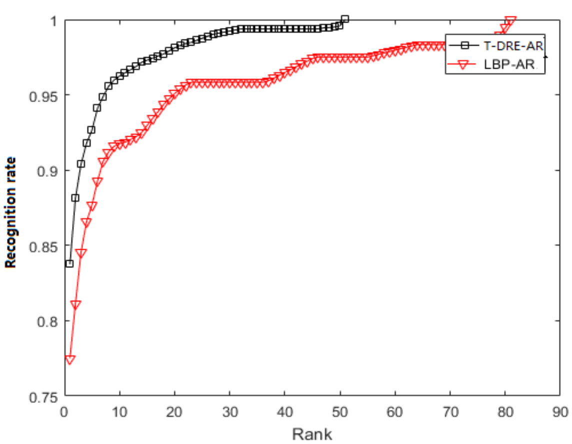

Figure 9. Recognition rate comparison of LBP and T-DRE in AR database

To further shown the experimental results more clearly, the comparison of recognition rate is shown in table1.

4. CONCLUSION

In this paper, a transfer learning based deep feature extraction framework (T-DFE) is presented for human identification using face images. The main purpose is to conquer the disadvantage of traditional CNN on small sample task. Firstly, a simply convolutional neural network is trained using several augmented database with abundant sample numbers of each subjects. Then the chosen pre-trained deep network is transfered to conduct recognition of database with small sample of each person. To further simplify the proposed deep network and enhance its flexibility, the softmax layer is removed when conducting transfer learning. The classification task is conducted by Euclidean Distance. Abundant comparison experiments on ORL database, IMM database and AR database have shown the effectiveness of the presented method.

Table 1

Database	Algorithm	Accuracy
ORL face database	T-DRE	97.25%
	LBP	78.25%
IMM face database	T-DRE	97.6667%
	LBP	81.63%
AR face database	T-DRE	83.75%
	LBP	77.5%

5. ACKNOWLEDGEMENTS

This work is supported by the National Key R&D Program of China under Grand 2017YFB0802300, Research Project of Beijing Municipal Education Commission (Grand No. KM201810009005), the North China University of Technology “YuYou” Talents support Program, Beijing Natural Science Foundation (4162022), High Innovation Program of Beijing(2015000026833ZK04).

6. REFERENCE

[1] Zhang D, Lu G, 3D Biometrics.2013. *System and Application*.

[2] Sharif M, Bhagavatula S, Bauer L, et al. Accessorize to a Crime: Real and Stealthy Attacks on State-of-the-Art Face Recognition// ACM Sigsac Conference on Computer and Communications Security. ACM, 2016:1528-1540.

[3] OM Parkhi, A Vedaldi, A Zisserman.2014. Deep Face Recognition. *British Machine Vision Conference*.2015(41).1-12.

[4] Zhao W, Chellappa R, Rosenfeld A, et al.2003. Face Recognition: A Literature Survey. *ACM Computing Surveys* ,35(4),399-458.

[5] Guo Y, Zhang L, Hu Y, et al. MS-Celeb-1M: A Dataset and Benchmark for Large-Scale Face Recognition. 2016:87-102. 6.

[6] Geng M, Wang Y, Xiang T, et al. 2016. Deep Transfer Learning for Person Re-identification.

[7] M Oravec. 2014. Feature extraction and classification by machine learning methods for biometric recognition of face and iris. *Elmar*.

[8] Masi I, Trần A T, Hassner T, et al. Do We Really Need to Collect Millions of Faces for Effective Face Recognition. 2016:579-596.