Face Recognition Algorithm Based On Convolutional Neural Network

Shiyuan Xu 1*, Shicheng Xu 2, Zhaojun Wang 3, Chengyan Pu 4, Qingchuan Meng 5

1 North China University of Technology, BeiJing, BeiJing, 100144, China
2 Harbin Institute of Technology, WeiHai, ShanDong, 264200, China
3 Beijing City University, BeiJing, BeiJing,101399, China
4 University of Electronic Science and Technology of China, ChengDu, SiChuan, 611731,China
5 HeiLongJiang Province Cyberspace Research Center, Harbin, HeiLongJiang, 150090, China

*Corresponding author’s e-mail: kingge666@163.com

Abstract: In response to the needs of modern enterprises for intelligent security of face authentication, a face recognition algorithm based on deep convolutional neural network is introduced in this paper. In order to improve the accuracy of face recognition, the design of 11 convolutional layers and 4 pooling layers. The network uses the standard face data set CASIA-WebFace for training, and the face recognition accuracy on the LTW database can reach more than 97.8%. Based on this deep learning network, a face recognition management system was designed. The system realized 1:1 face authentication and 1:N face recognition. After applying the convolutional neural network algorithm, the face search time is significantly reduced, indicating that the algorithm has higher efficiency and practicability.

1. Introduction
In recent years, with the improvement of computer hardware and the increase of data volume, face recognition has become one of the research hotspots in the field of computer vision. Face recognition is a comprehensive biometric technology, which includes image acquisition, face detection, face alignment, feature extraction, and face database comparison. Scenes. Research on face recognition technology at home and abroad began in the 1960s. Since 2014, thanks to the massive increase in the amount of network data, "big data + deep learning" has become an important way to study face recognition. In this context, the research on neural networks, training techniques, and data sets related to deep learning has also made breakthrough progress. However, due to the similarity between different faces and the variability of the same face, current face recognition is difficult. Aiming at this problem, this paper has conducted in-depth research on the structure and training methods of neural networks in deep learning theory to improve its recognition efficiency under massive data [1-4].

This paper designs a face recognition management system based on convolutional neural networks (CNN). This system can realize face authentication and local self-built database search [5-6]. The content of this article includes system overview and system implementation.
2. System overview
This paper designs the system based on software engineering theory. The first step in system design is to perform system requirements according to the purpose of the design, decompose the goals, and divide them into functional modules. The purpose of the face recognition management system designed in this paper is to use the unique biological information of the face to complete the student identity authentication.

According to the purpose that the system needs to achieve, two recognition technologies are included in the field of face recognition technology. The first is a 1:1 face authentication technology: the system needs to determine whether any two given pictures belong to the same picture. In this mode, the essence of face recognition is a binary classification problem. There are only two results: yes or no; the second is a 1: N face recognition technology: that is, when a face picture is given, the face closest to the face needs to be retrieved from the database To learn their identity. According to the analysis in this section, combined with the practical application of the two types of identification technology, the system can be divided into functional architectures as shown in Figure 1.

As shown in Figure 1, the system is deployed using a Client-Server (C/S) architecture. The front end of the system is divided into two functional modules: user interaction and result display. The back end is divided into four functional modules: security check, face detection, and face certification, template rendering. The specific functions of each module are as follows:

User interaction: In this module, the user completes related interactive operations with the system. This module is the user interface of the system. This module performs operations such as uploading face images and querying face authentication results.

Results display: In this module, the face discrimination result obtained from the backend is returned for users to view.

Security check: As the system uses the method of web crawler to expand the local face image database, it is necessary to perform security verification of new images when the database is expanded. After the security check, the security of the image source can be guaranteed.

Face detection: When the system uploads the user’s image, it needs to determine whether the image transmitted by the user contains the image information of the face. Therefore, a face detection module needs to be introduced. After the module determines that the image contains face information, the system will further process the image, such as alignment and normalization.
Face certification: This module is the core module of the system. This module calls the designed deep learning algorithm to complete the 1:1 authentication of the face. In addition, this module can also be used for face search, that is, given in the local database and the user enters the most similar results.

Template rendering: The design of this module is mainly set from the perspective of system implementation. In the design of the user interface, the system uses Jinja2 of Flask in Python. Jinja2 needs to define the base template for hierarchical design when implementing the interface.

3. Algorithm design
In order to complete the design of the face authentication module in the system function module, it is necessary to use the corresponding face recognition algorithm. In this paper, a recognition algorithm based on deep learning theory is used [5-6]. Deep learning (DeepLearning) refers to the establishment of multilayer artificial neural networks to guide computers to complete artificial intelligence in corresponding fields. In the field of computer vision, face recognition and authentication can be achieved through deep learning. The deep neural network used in this paper is Convolutional Neural Network (CNN). In CNN, it can be divided into four different categories of levels: convolutional layer, pooling layer, activation function layer, and classification layer.

3.1. Convolutional layer
In the convolution layer, the image is regarded as a signal, and the correlation between the image signal and the convolution kernel is calculated using a convolution operation. The calculation method is as follows:

\[
X(x, y) * W(x, y) = \sum_{s=-h}^{h} \sum_{t=-h}^{h} W(s, t) X(x-s, y-t)
\]  

(1)

Where \(X\) is the two-dimensional signal of the input image and \(W\) is the convolution kernel used. During the convolution process, \(W\) moves one pixel at a time on \(X\) to get the final matrix \(S\). In the same convolution layer, it usually contains \(N\) different convolution kernels, and different convolution kernels correspond to different feature outputs. In the end, this layer will get \(N\) feature maps as input to the next layer of the entire CNN network.

3.2. Pooling layer
The pooling layer is a process of downsampling. Aggregation is performed by matrix elements adjacent to the pixel. Common aggregation methods include the regional average or maximum. When using max for pooling:

\[
x_j^l = f(\beta^l \max_{i \in M_j} x_i^{l-1} B^l)
\]  

(2)

3.3. Activation function layer
Due to the transmission through a multi-layer network, the output distribution characteristics obtained by the network are different and need to be normalized. which is:

\[y = f(\sum w_i x_i + b)\]  

(3)

Among them, the commonly used \(f\) are Sigmoid function, Tanh function, and Relu function.

3.4. Classification layer
In this paper, when classifying, this problem is regarded as a regression problem of probability, and the Softmax model is used. When the number of input data is \(x\) and there are \(n\) categories, the probability distribution of category \(y=i\) can be expressed by equation (4):
\[ p(y = i \mid x; \theta) = \frac{e^{T_i x}}{\sum_{j=1}^{k} e^{T_j x}} \]  \hspace{1cm} (4)

Obviously, the probability of \( X \) under different labels can be obtained by equation (4). The label with the highest probability is its classification.

4. System implementation

4.1. CNN network design

In the system's face certification module, an external algorithm interface needs to be designed. This article designs the deep learning network structure shown in Figure 2.

![Deep learning network](image)

**Figure 2. Deep learning network**

In this network, there are 11 convolutional layers, 4 pooling layers, and 3 fully connected layers. Among them, the activation function layer selects the Relu function, and the classification layer selects the Softmax model.

4.2. Algorithm simulation

When training the network, this article uses the CASIA-WebFace standard face library. The inventory is at 494414 face images of 10575 people. Before training, the face image is pre-processed. The pre-processing process includes the alignment and normalization of human faces, while converting the image into a 100x100x1 grayscale image. The training process includes two processes, forward propagation and backward propagation. After training, multiple training models can be obtained. Through multiple tests, the network with the highest accuracy is taken as the final face recognition model.

When measuring network performance, this article uses the LTW standard face database. In this library, 13233 face images of 5749 people are included. According to the LTW test standard, the composition of the database is shown in Table 1. This article only uses the View2 part in LTW. View2 is divided into 10 subsets, each of which contains 600 human face samples. There are 300 positive samples and 300 negative samples. For network performance evaluation, a 10-fold cross-validation method is used. For each test, 1 test set in 10 subsets is selected, and the remaining 9 are used as the training set. At this time, the average accuracy is defined as follows:

\[ \text{accuracy}_{\text{average}} = \frac{1}{10} \sum_{i=1}^{10} \frac{c_i}{N_i} \]  \hspace{1cm} (5)

The test results of this paper are compared with existing deep networks, as shown in Table 2. It can be seen that the accuracy rate of the face recognition method proposed in this paper is significantly
higher than other deep networks, and the accuracy rate is 98.78%. At the same time, the retrieval and recognition time of the face image of this method only takes 0.525s, which has also been significantly reduced compared with other deep networks.

Table 1. LTW library composition

| View 1          | View 2          |
|-----------------|-----------------|
| Application     | Model selection and development | Algorithm test |
| Number          | -               | 6000            |

Table 2. Performance of different deep learning algorithms on LTW

| Deep learning algorithm | Accuracy rate | Number of training samples | Retrieval and identification time |
|------------------------|---------------|----------------------------|----------------------------------|
| DeepFace               | 94.35%        | 700000                     | 0.898s                           |
| DeepID                 | 95.24%        | 500000                     | 0.736s                           |
| CNN                    | 98.78%        | 494414                     | 0.525s                           |

5. Conclusion

According to the security requirements of face authentication, this paper designs a face recognition management system. In the process of system design, the method first studied the theory of deep learning and improved the CNN network. In this paper, the design and structure of the network are discussed in detail. After testing, the accuracy of face recognition on the standard data set is improved to 98.78%. This network can be integrated as an algorithm module and connected to any face recognition management system. At the end of this article, the trained neural network is applied to a 1:N face search system. After testing, the time for retrieving and recognizing a face image is much lower than other deep networks, and the recognition accuracy and efficiency are greatly improved.

References

[1] Zhang, L., Liu, J., Zhang, B., et al. (2019) Deep Cascade Model-Based Face Recognition: When Deep-Layered Learning Meets Small Data. IEEE Transactions on Image Processing., 29: 1016–1029.

[2] Lu, J.W., Liong, V.E, Wang, G., et al. (2015) Joint Feature Learning for Face Recognition. IEEE Transactions on Information Forensics and Security., 10: 1317–1383.

[3] Gao, S.H., Zhang, Y.T., Jia, K., et al. (2015) Single Sample Face Recognition via Learning Deep Supervised Autoencoders. IEEE Transactions on Information Forensics and Security., 10: 2108–2118.

[4] Lu, J.W., Wang, G., Zhou, J. (2017) Simultaneous Feature and Dictionary Learning for Image Set Based Face Recognition. IEEE Transactions on Image Processing., 26: 4042–4054.

[5] He, R., Wu, X., Sun, Z.A., et al. (2019) Wasserstein CNN: Learning Invariant Features for NIR-VIS Face Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence., 41: 1761–1773.

[6] Kang, S., Lee, J., Boog, K., et al. (2018) Low-Power Scalable 3-D Face Frontalization Processor for CNN-Based Face Recognition in Mobile Devices. IEEE Journal on Emerging and Selected Topics in Circuits and Systems., 8: 873–883.