Research on Face Recognition Based on Improved Convolutional Neural Network Using Raspberry Pi

Lingyi Ning, Zhikai Huang1*, Xianchang Xi and Yang Zhang

1College of Information and Engineering, Nanchang Institute of Technology No. 289 Tianxiang Avenue, Changdong Town, Nanchang County, Nanchang, China
E-mail: huangzhik2001@163.com

Abstract. In this paper, a face recognition system has been designed using Raspberry Pi based on convolution neural network technology. Convolution neural network that using OpenCV (open source computer vision library) to realize real-time face recognition and feature extraction. The difficult and important problem in deep learning network is that it could be over fitting very easily. In order to alleviate the over fitting phenomenon and improve the accuracy of face recognition, the dropout technology of convolutional neural network had been employed to improve the recognition accuracy and ensure the robustness of face recognition. The results of experiments show that our methods has the advantages in the computer time and recognition rate for the low construction cost, simple structure and strong portability. The accuracy of face detection for static and dynamic images using Raspberry Pi is above 85% in our testament.

1. Introduction
Artificial intelligence technology has been widely used in our life, especially in the field of computer vision and biological feature extraction [1]. Combined with machine learning, artificial intelligence and other technologies, it can be widely used in finance, security industry and face attendance and so on [2]. The focus of face recognition system is it’s recognition speed and accuracy. As a flexible hardware device, the Raspberry Pi has good image capturing functions. The method of convolutional neural network is used to realize face recognition on the Raspberry Pi. Compared with traditional face recognition methods, deep learning can learn the representation features of faces from massive amounts of data [3]. This paper intends to design and develop a face recognition system on the Raspberry Pi, based on the improving convolutional neural network algorithm [4]. This method can quickly adapt to different face analysis tasks and has the advantages of fast recognition speed and higher accuracy.

2. Design of Face Recognition System Based on CNN
The face recognition system based on the Raspberry Pi microcomputer uses OpenCV to detect and recognize faces [5]. The system uses Raspberry Pi and HD camera as the main hardware to build OpenCV face detection environment. The improved convolutional neural network algorithm is combined with image processing function in OpenCV to realize the face detection in the video images.

As shown in figure 1(a) and (b), Raspberry Pi is used as the server, using multi-threaded design, through a main thread to start image acquisition, image processing, steering gear control, communication and other sub-threads. Each thread realizes interaction and data transmission by sharing global variables. This method can make the image acquisition smooth.
A personal computer is used as client, and a window is used for development. Because the performance of desktop is generally higher calculate ability, the training, recognition and database functions are all placed on the PC side. The client also adopts multi-threaded design. The schematic diagram of the client software structure is shown in figure 2. Figure 2 (a) software structure function diagram, figure 2 (b) software structure module diagram.

Design and develop GUI interactive in desktop with pyqt5. The UI thread is responsible for interacting with the user, displaying both images and the server-side data information. The interaction page has been shown in figure 3. The information entry interface is shown in figure 3(a). It is used for the information input of the newly-recognized object and the corresponding facial image collection. The corresponding facial image acquisition is shown in figure 3(b). The input information will be stored in the database and mapped with the training data one by one. The data management window is used for look up, update, and deletes management of existing information, which can easily maintain and update existing personnel information.
3. Convolutional Neural Network Model

Convolutional Neural Network is evolved from artificial neural network, and each layer contains many two-dimensional planes [6]. Figure 4 shows the network model of CNN. In general, layer C represents feature extraction, and layer S represents feature mapping. The feature extraction layer (C layer) must be followed by the calculation layer (S layer). The calculation layer will calculate the local average the features of the feature extraction layer locally, and then extract the features again. In this way, the network can produce anti-distortion capabilities for the sample and make identification more robust.

The system can be nonlinear by introducing the activation function Sigmoid [7]. Each layer of a Convolutional Neural Network basically uses an activation function. The sigmoid function curve is S-shaped and is a centrally symmetrical image. The centre of symmetry is (0,0.5), and the value range is (0,1). The curve is not at the origin, which is similar to having an offset term, so when it acts as an activation function; it has its own offset. The function and its derivative are:

\[ f(x) = \frac{1}{1 + \exp(-x)} \]  \hspace{1cm} (1)

\[ f'(x) = \frac{e^{-x}}{(1 + e^{-x})^2} \]  \hspace{1cm} (2)

The most serious problem of deep learning networks is over fitting, and it is also a common problem. If the sample data has been trained to fit in face detection, then face recognition is failed [8]. Dropout technology can effectively alleviate the over-fitting phenomenon, so the dropout layer is added to the CNN training stage.

The core of dropout technology is to average combine different sparse models. Pick some neurons at random and let them work together. The joint adaptability between nodes is reduced and the generalization is enhanced [9]. Figure 5 is a network model with and without dropout technology.
From the figure, it can be seen that the neurons between the layers of the fully connected network are continuously connected in sequence. However with the dropout technology, the neuron connections between the layers are randomly sparse.

![Network Diagrams](image)

**Figure 5.** Network model (a) Without dropout technology (b) With dropout technology

It can be seen from figure 6 that the calculation formulas of the network neural node without Dropout technology are shown in (3) and (4). In the formula W represents the network weight, b represents the offset value, z is the output value, and f is the excitation function.

\[
z^{(l+1)}_i = w^{(l+1)}_i y^{(l)} + b^{(l+1)}_i \tag{3}
\]

\[
y^{(l+1)}_i = f(z^{(l+1)}_i) \tag{4}
\]

The calculation formula of the network node with dropout technology are as follows:

\[
r^{(l)}_j = Bernoulli(p) \tag{5}
\]

\[
\overline{y}^{(l)} = r^{(l)} \ast y^{(l)} \tag{6}
\]

\[
z^{(l+1)}_i = w^{(l+1)}_i \overline{y}^{(l)} + b^{(l+1)}_i \tag{7}
\]

\[
y^{(l+1)}_i = f(z^{(l+1)}_i) \tag{8}
\]

Comparing the above formulas, after adding the dropout technology, the upper and lower neurons will use the Bernoulli random distribution function to determine whether they are connected [10]. Connected nodes may be different each time, and a diverse sparse network is formed. After verification, when the dropout value is 0.5, the most network models can be randomly generated, and the best network effect can be obtained by training.

The final classification of the CNN network depends on the soft max classification function, and the classification result is determined by the probability. When the input \( x \) is given, the expression of
the classification category \( j \) is equation (9), \( y^{(i)} \) represents the output category, \( \theta \) represents the connection weight, and the output value of \( e^x \) is increases exponentially. If (9) formula \( p > 0.5 \), then the input is expressed as type \( j \), and the function can be decomposed as shown in (10):

\[
p(y^{(i)} = j | x^{(i)}; \theta) = \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^{n} e^{\theta_l^T x^{(i)}}}
\]

(9)

\[
e^\theta_j^T x^{(i)}
\]

(10)

In (10), \( \theta^k \) represents the connection weight of category \( k \). \( \theta_j \) represents the connection weight of category \( j \). The Dropout layer ranks the output values from large to small and then removes the less effect to the classification of network node. The improved algorithm is integrated into the CNN network. In the dropout module, the sparsely allocation calculation sub module is connected, and the two levels are connected randomly through the redefined sparsely.

4. Test and Result Analysis

4.1. System Test

The whole face recognition workflow of the system is to load the user’s face image into the database, and name the users, then retrain. The system can register multiple new users according to the actual needs. Face recognition system was tested online in Windows 10 environment, as shown in figure 7, 7(a) is a single online test, 7(b) is a multi-person online test.

(a)  
(b)

Figure 7. Face recognition system test online (a) Single online test; (b) Multi-person online test.

4.2. Static Face Detection

Using the above system, 400 images from the ORL and 400 images from the WIDER FACE database were selected randomly as samples for static face detection test. The experimental results are shown in table 1, and the detection rates are 92.25% and 90.75% respectively.

| Image type  | Experiments /times | Correctly detect /pieces | Missed detect /pieces | False detect /pieces | Detection rate/% | Missed detection rate/% | False detection rate/% |
|-------------|--------------------|--------------------------|-----------------------|----------------------|------------------|------------------------|-----------------------|
| ORL         | 400                | 369                      | 31                    | 0                    | 92.25            | 7.75                   | 0                     |
| WIDER FACE  | 400                | 363                      | 29                    | 8                    | 90.75            | 7.23                   | 2.02                  |

Table 1. Database picture sample test results.
Face images are obtained from the camera by the detection system, and the single person and multi person images are detected respectively. The experimental results are shown in table 2, and the detection rates are 90% and 85% respectively.

| Image type  | Experiments /times | Correctly detect /pieces | Missed detect /pieces | False detect /pieces | Detection rate/% | Missed detection rate/% | False detection rate/% |
|-------------|------------------|-------------------------|----------------------|----------------------|-----------------|------------------------|------------------------|
| Single image| 40               | 36                      | 4                    | 0                    | 90              | 10                     | 0                      |
| Multi-image | 40               | 34                      | 4                    | 2                    | 85              | 10                     | 5                      |

4.3. Dynamic Face Detection

The dynamic video stream is obtained through the camera. Calling the modules in the OpenCV, face detection is carried out on each frame of the video, which is divided into two groups of data: single person image and multi person image. The interval of each experiment was set as 10s. If the face detected and labelled correctly, it is regarded as successful; if the face not labelled in areas where the face has, it is regarded as missed; if the face labelled in areas where has no face, it is regarded as false detection. The results of the detection experiment are shown in table 3 below. The face detection rate of single-person images is 87.5%, and the face detection rate of multi-person images is 85%.

| Image type  | Experiments /times | Correctly detect /pieces | Missed detect /pieces | False detect /pieces | Detection rate/% | Missed detection rate/% | False detection rate/% |
|-------------|------------------|-------------------------|----------------------|----------------------|-----------------|------------------------|------------------------|
| Single image| 40               | 35                      | 4                    | 1                    | 87.5            | 10                     | 2.5                    |
| Multi-image | 40               | 34                      | 4                    | 2                    | 85              | 10                     | 5                      |

4.4. Result Analysis

Experiments are carried out on the improved CNN structure. Five convolution layers and two fully connected layers are selected in the network. Each convolution layer is followed by a pooling layer. The improved dropout layer is added to the sixth layer. The same face data set as above is selected for training and recognition with the face image collected in advance. Compared with the unimproved CNN, the accuracy rate and model loss rate of the experimental results are shown in figure 8, figure 9:

**Figure 8.** CNN model (a) Accuracy rate; (b) Loss rate.
From the comparison of the above figures, training recognition on the unimproved convolutional neural network, accuracy rate tends to about 82%, and the loss rate tends to 0.4. However experiments on the improved CNN structure, the accuracy rate tends to about 88% and the loss rate is relatively low and tends to be 0.2. It shows that the improved CNN model has good applicability in this face recognition task.

5. Conclusion
The system is compiled with Python language, combined with OpenCV face detector and improved CNN algorithm to complete online face recognition. Compared with the traditional face analysis method, the improved CNN algorithm does not need to deliberately design face features, but automatically learns face representation from massive data through neural network. This not only simplifies the complex feature extraction, but also can learn some hidden rules and representations in the face image through neural network. This algorithm has the advantages of low construction cost, simple structure and strong portability. The face detection accuracy of static and dynamic images is more than 85%, which not only improves the recognition accuracy, but also ensures the robustness of the algorithm.

Acknowledgement
This project was supported by National Key Research and Development Plan of China (No.2019YFB1704502), the National Natural Science Foundation of China (Grant No. 61472173), the grants from the Educational Commission of Jiangxi province of China, No. GJJ151134, the Innovation Fund Designated for Graduate Students of Jiangxi Province, No. YC2020-S633.

References
[1] Taigman Y M, Ranzato M A, et al. 2019 Deepface: Closing the gap to human-level performance in face verification Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition: 1701- 1708.
[2] Venkatesh B, Harshav K, Santosh J, et al. 2018 Door security using face detection and raspberry Pi Proceedings of Materials Science and Engineering 331: 012011-012018.
[3] VAPNIK V, CORTES C 2020 Support-vector networks Machine Learning 20(3): 273-297.
[4] Wen Y, Zhang K, Li Z, et al. 2019 A discriminative feature learning approach for deep face recognition European Conference on Computer Vision Spring, Cham: 499-515.
[5] Huang G, Liu Z, Van Der Maaten L, et al. 2020 Densely connected convolutional networks Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition: 4700-4708.
[6] Lv W H and Lei J Y 2020 Deep learning development review Proceedings of the 2020 IEEE 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE): 171–174.
[7] Mohammed I M, Al-Dabagh M, Ahmad M I, et al. 2019 Face recognition using PCA implemented on raspberry Pi Proceedings of the 11th National Technical Seminar on Unmanned System Technology: 873-889.
[8] Sajjad M, Nasir M, Muhammad K, et al. 2017 Raspberry Pi assisted face recognition framework for enhanced law-enforcement services in smart cities *Future Generation Computer Systems*: 995-1007.

[9] Sajjad M, et al. 2017 Raspberry Pi assisted face recognition framework for enhanced law-enforcement services in smart cities *Future Generation Computer Systems*: 995-1007.

[10] Faisal F and Hossain S A 2019 Smart security system using face recognition on raspberry pi *Proceedings of the 13th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)* IEEE.