

# Research on Human Eye Iris Recognition Based on Inception v3

Yuewen Tang\*, Xiangkui Li, Yulian Jiang

Southwest Minzu University, ChengDu, China  
Email: tyw663852577@qq.com

**Abstract.** Iris recognition has certain application prospects in access control systems, financial fields, and airport security. Before the appearance of the deep learning method, the iris recognition method relies on artificial method for feature extraction. This method will consume a lot of time and affect the efficiency of recognition. In this recognition study, we adopt the method of deep learning. On the basis of a large amount of image data and the guarantee of computer computing ability, in order to exceed the iris recognition accuracy of previous recognition methods. In this paper, the improved Inception v3 model is used to classify and recognize the iris database provided by the Chinese Academy of Sciences in the Tensorflow framework. Finally, the accuracy of image recognition can reach 98% through the training of the model.

**Keywords:** Inception v3, Tensorflow, iris recognition.

## 1 Introduction

Certain vital signs of humans are unique, such as fingerprints, tongue lines and iris. These features are being used in fingerprint unlocking, palmprint unlocking and iris unlocking, and are gradually being used in access control systems, finance and airport security[1]. Iris recognition, which is based on the iris in the eye, is more secure than other biometric recognition.

The traditional iris recognition process is complex, including iris image quality assessment, localization, normalization, mask estimation, feature extraction, matching and rough classification. In addition, the extraction of iris features is usually artificial[2]. This traditional technology, step by step, takes a long time to identify.

With the rise of artificial intelligence, the ability of machine learning to deal with massive data is greatly enhanced, it can train large-scale data sets well, and the model has the ability of autonomous learning. In the field of machine learning, deep learning is an important branch of research and development. One of the core problems of deep learning is to automatically combine simple features into more complex features and use these combinatorial features to solve problems. Figure 1 shows the difference between traditional machine learning algorithm and deep learning algorithm.

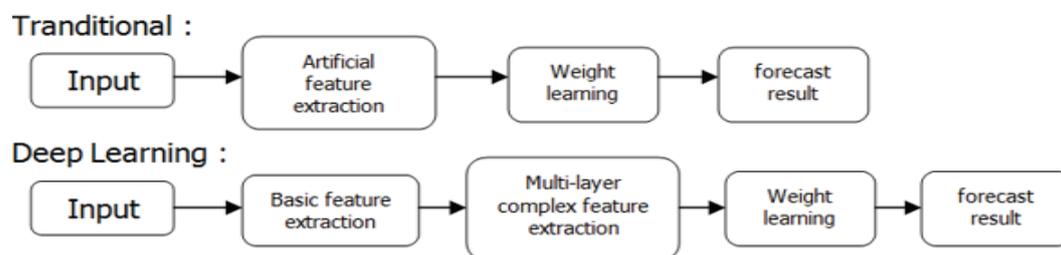


Figure 1. Comparison of traditional machine learning and deep learning algorithm steps.

As can be seen from Figure.1, the depth learning algorithm can avoid the complicated work of the artificial extraction features so as to have more independent learning ability, and the extraction of the features is simplified to propagation, so that the recognition accuracy is higher for the identified object.

In this paper, the deep learning Inception v3 model is applied to iris recognition. The iris database provided by the Chinese Academy of Sciences is used to train the model and make the model self-study, and finally identify the iris in the iris database of the human eye. Based on the deep learning method, the quality evaluation, localization, normalization, mask estimation, feature extraction and matching steps of the iris image in the traditional iris recognition can be completed by the neural network model self-training. Thus, while ensuring the accuracy of recognition, the speed of recognition is improved.

## 2 Tensorflow Framework

Tensorflow is a framework platform that can support various deep learning algorithms. It has the advantages of flexible expression, fast implementation of various algorithms, high execution performance, distributed scalability, cross-platform portability, experimental reproducibility, support for rapid production, and the model can be deployed at any time[3]. It can support a lot of computer languages, such as Python, C, Go, Java, and so on[4]. The most intuitive advantage of the Tensorflow framework and the Matlab is time saving. The application of the large-scale deep learning model established by Tensorflow is very wide, including speech recognition, natural language processing, computer vision, robot control, information extraction, drug development, molecular motion prediction, etc.

## 3 Inception v3 Model

Before the emergence of the Inception model, most of the breakthroughs in the mainstream network structure were to increase the number of layers to make the network deeper, and at the same time to increase the number of neurons connected to each layer in order to make the network wider. However, the disadvantages of these operations are as follows: when the training data set is limited, too many parameters will lead to over-fitting of the model, and the complexity of calculation will increase greatly, so it will be difficult to design it. Even when the number of layers of the network increases, the gradient transfer becomes easier and easier to disappear. The Inception model came into being to overcome these defects. The structure of the model is as shown in Figure 2.

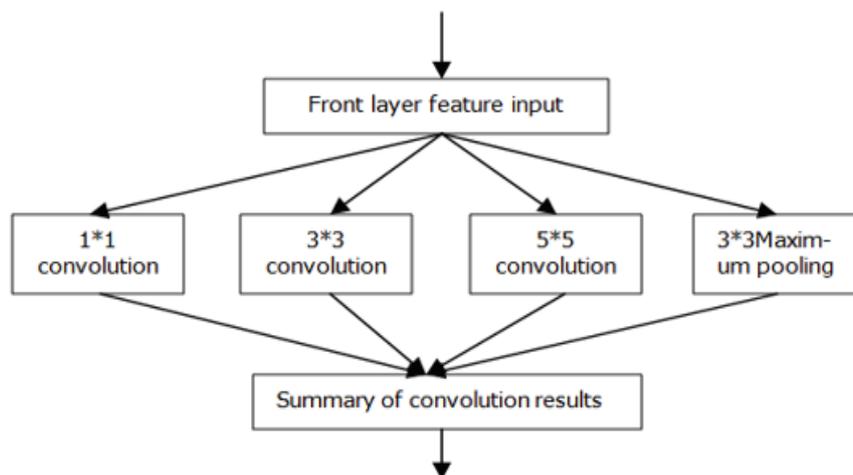


Figure 2. Inception model structure.

The Inception v3 model is improved by the Inception v1 and Inception v2 models. The Inception v1 model has small network depth and width, and there are obvious deficiencies in the recognition accuracy of the model. The Inception v2 model improves the recognition accuracy. In the recognition process, the Batch Normalization method is proposed to replace the Dropout and LRN methods of the first generation model. The Inception v2 model replaces the 5×5 convolution of the Inception v1 model with two layers of 3×3 convolution, and the computational time saved is used to increase the number of filters, due to the limited depth and width of the network, the recognition accuracy still fails to reach the

ideal peak, and the recognition speed needs to be improved. The Inception v3 model introduces the Factorization method, which splits a large two-dimensional convolution into two smaller one-dimensional convolutions. In the Inception v3 model, for a  $5 \times 5$  convolution kernel convolution region, a  $5 \times 5$  region can be convolved with a  $3 \times 3$  convolution kernel, and then a  $3 \times 3$  kernel convolution again then we will get a data. The effect is the same as the  $5 \times 5$  convolution. This reduces the number of convolution calculations and makes the recognition of images faster. The Inception v3 model structure is shown in Figure 3.

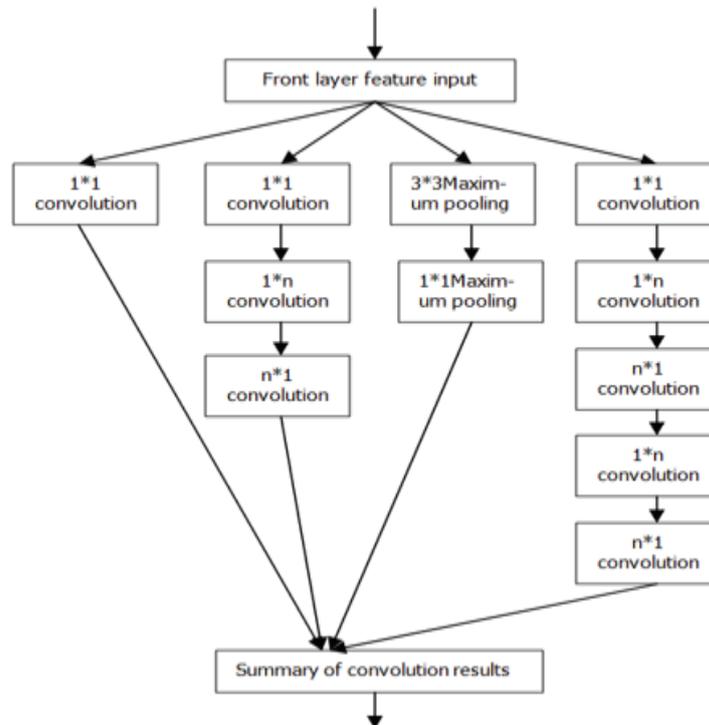


Figure 3. Inception v3 model structure.

Inception v3 model has the function of efficient and accurate image classification. If it is transplanted to the application end, it will also have a good market prospect. Based on the above advantages, the Inception v3 model is used to classify the iris.

## 4 Tensorboard

Tensorboard is the visualization tool of Tensorflow[5]. It can visualize the running state of Tensorflow program through the log file output in the running process of Tensorflow program[6]. With this visualization tool, we can see the running structure of the Inception v3 model more specifically and clearly. Besides the traditional convolution layer and pool layer model, we can also observe that there are many mix layer in network, as shown in Figure 4.

The mix layer contains one convolution layer and three tower layers. The purpose of the mix layer is to increase the width of the model, while the number of network layers is to increase the depth of the model. The mix layer is called Inception. Of course, the designed inception can only deal with 1000 categories of images. If we want to use this model to train the less common images, we need to improve on the basis of them.

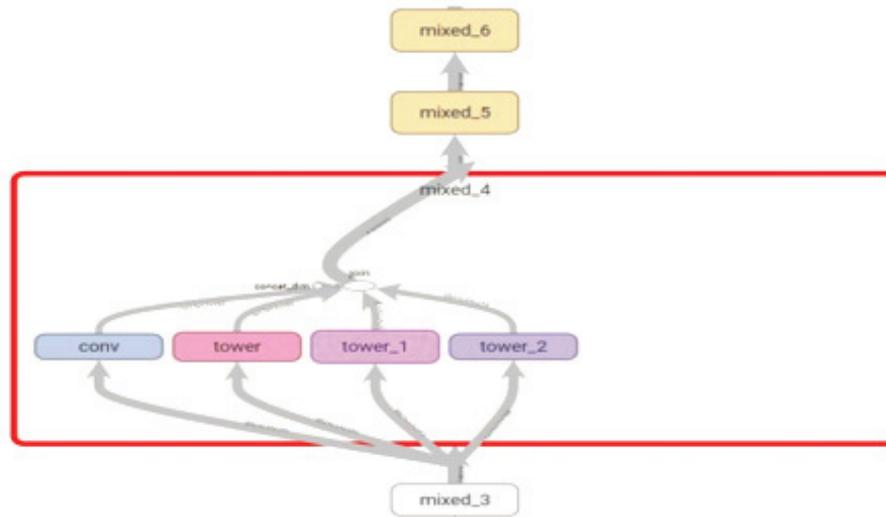


Figure 4. Expanded mix network layer.

### 5 Human Eye Iris Recognition Based on Improved Inception v3 Model

The efficient recognition of the Inception v3 model lies in its algorithmic structure, where data is transferred from bottom to top, and all data is fixed at all the network layers below the last pool layer. What needs to be changed is the network structure of the last pool layer. The whole network shares weights, so the speed of image recognition is faster. When there is input, the identified confidence level can be quickly obtained by adding a softmax function to the last pooling layer. In this paper, the improvement of the Inception v3 model is shown in Figure 5.

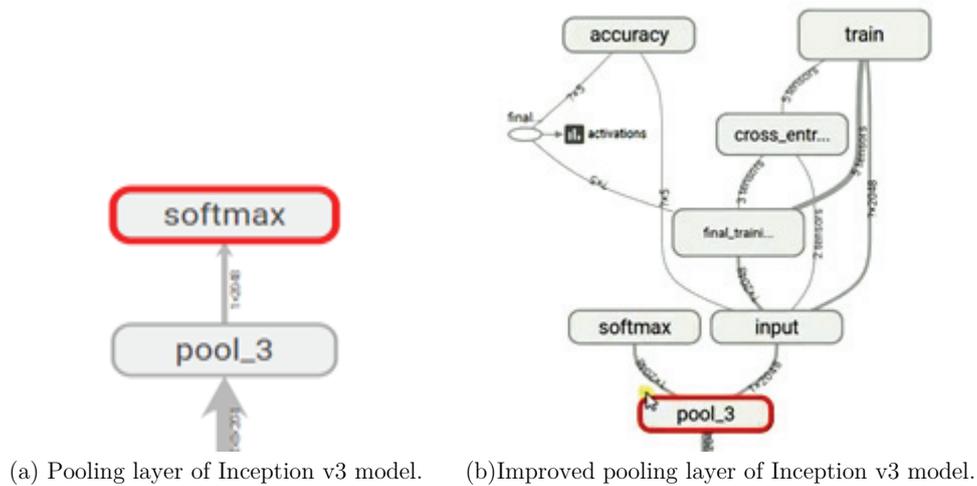
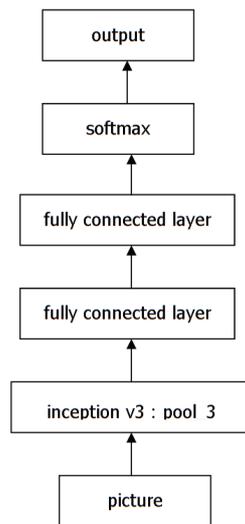


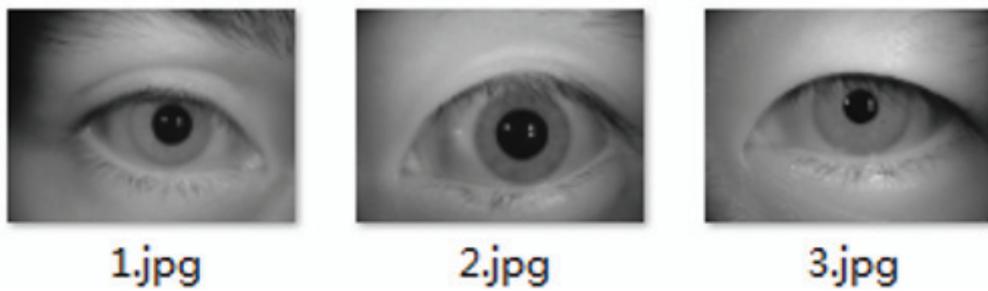
Figure 5. Improved schematic of the Inception v3 model.

As can be seen from Figure 5, the network we want to train adds an input layer behind the pooling layer, and the data set needs to be trained from this location and then input. In order to enable the network structure to extract more feature information from the picture, we added two fully connected layers after the pool\_3 layer to expand the network depth. The model after adding two fully connected layers is shown in Figure 6.



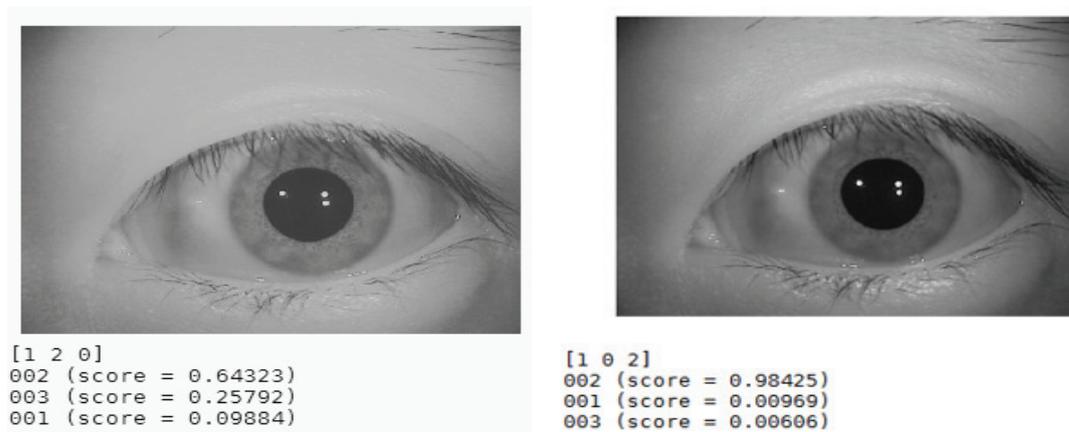
**Figure 6.** Network structure after change.

Using the modified model structure to identify the iris image, we took three different individual iris images to distinguish our recognition results, as shown in Figure 7.



**Figure 7.** Iris to be identified.

Identifying the image with the improved Inception V3 model will result in a confidence level, as shown in Figure 8.



**Figure 8.** Recognition and comparison of pictures before and after modification.

As can be seen from Figure 8, the model can accurately identify the image with a confidence level of 98%. This is because we have added two fully-connected layers after the model Pool\_3 layer, which deepens the network structure of the model and enables the network structure to acquire more feature information from the image to be identified, resulting in the identification of the model. Accuracy is improved. Table 1 compares the recognition accuracy of the proposed algorithm with the traditional algorithm.

**Table 1.** The accuracy of the algorithm and traditional algorithm identification.

Algorithm name	Recognition accuracy /%
Boles Algorithm	93.56
Gabor Filter	97.24
Inception v3 Model	64.32
Improved Inception v3 Model	98.42

From the experimental results in Table 1, we can see that compared with the traditional iris recognition method, the proposed algorithm based on Inception v3 model improves the recognition accuracy to a certain extent.

## 6 Conclusion

In this paper, the deep learning algorithm is used to solve the problem of low efficiency of traditional iris recognition methods. The Inception v3 model has the advantage of fast recognition of target images. However, the recognition accuracy is insufficient due to the small number of training samples, so the network structure is readjusted in this paper, and one full connection layer is decomposed into two full connection layers on the full connection layer of the original model. The improved Inception v3 model is applied to the target image recognition of the human eye iris database provided by the Chinese Academy of Sciences. The improved Inception v3 model has higher recognition accuracy than the Boles algorithm and the Gabor filter algorithm, and the recognition accuracy of the improved V3 model is higher than that of the Boles algorithm and the Gabor filter algorithm. The reason is that the deep learning model has a good self-learning ability and constantly enhances the recognition ability of the model to the target picture during the training process of the model. In the following work, it is hoped that the experimental results will be combined with the hardware to make a finished product for application.

**Acknowledgement.** Thanks for the careful guidance of the teacher and the help of the classmates. This work was funded by the operation fee of the innovative research project of Southwest Minzu University, project number: CX2018SZ93.

## References

1. Deng L, Yu D. Deep learning: methods and applications[J]. Foundations and Trends in Signal Processing, 2014, 7(3-4):197-387.
2. Schmidhuber J. Deep learning in neural networks: An overview[J]. Neural networks, 2015, 61:85 — 117.
3. Li H. Deep learning for natural language processing: advantages and challenges[J]. National Science Review, 2017, 51(1):24-26.
4. Harmon L D, Khan M K, Lasch R, et al. Machine identification of human faces[J]. Pattern Recognition, 1981, 13(2):97-110.
5. Ding C, Tao D. Trunk-branch ensemble convolutional neural networks for video-based face recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 40(4):1002-1014.
6. Wang F Y, Zhang J J, Zheng X, et al. Where does AlphaGo go: From church-turing thesis to AlphaGo thesis and beyond[J]. IEEE/CAA Journal of Automatic Sinica, 2016, 3(2):113-120.