Facial detection and classification

Chao Li
Nanjing Agricultural University

Fan Zhang
Hunan University

Yanwen Huang
Hunan University

August 5, 2018
Preliminary

Abstract

Vision is a major way of human perception to the outside world, because of human visual properties, the research related to human , people can easily distinguish between two faces. With the rapid development of computer technology and the broad application in video monitoring system in the city, people began to think about how to make computer like people quickly find and recognize faces, help us to handle large amounts of video information,and solve the limitation of the traditional identity authentication system, enhancing the working efficiency of the related areas.

This paper constructs a set of face detection and recognition system based on real-time video, which mainly includes real-time detection and recognition of human face. In this paper, the source code of opencv and Japanese programmers is used to train their face recognition model by using the keras deep learning library, and the study, summary and in-depth discussion are carried out. At the same time, in order to verify the efficiency and accuracy of other deep learning libraries, Theano was also used to train the face recognition model using CNN - convolution neural network. At the same time, the project has been restored to the project, and some of the running bugs have been modified to improve the running smoothness of the program, improve the usability, and add the input name and other functions.

At the end of this paper, the author summarizes the work achievements and innovations of the system. Meanwhile, some shortcomings of the system are explained, and the future development is prospected.

1. Introduction

With the rapid development of modern information technology, the technology of identity authentication has shifted to the biological feature level.

Modern biological recognition technology is mainly through computer and technology closely, to identify a person using the inherent human physiological characteristics and behavioral characteristics of. Face recognition refers to the distribution of facial features and profiles of human faces, which vary from person to person and are innate.

Face is a set of patterns that contain rich information. It is one of the main symbols of human dialectics and recognition. It is also one of the objects of vision and interest in image and video. Compared with other biological features such as fingerprints, iris and voice, face recognition is more direct and friendly. It can achieve better recognition results without disturbing people's normal behavior.

We want to identify particular people in real-time (e.g., in a security monitoring system, location tracking system, etc.), or we want to allow access to a group of people and deny access to all others (e.g., access to a building, computer, etc.) [1]. Multiple images per person are often available for training and real-time recognition is required.[3]

In this paper, we are primarily interested in the case above. We are interested in recognition with varying facial detail, expression, pose, etc. We are interested in rapid classification and hence we do not assume that time is available for extensive preprocessing and normalization. Good algorithms for locating faces in images can be found in [5], [2], and [4]

The remainder of this paper is organized as follows. The data we used is presented in Part 2 and related work with this and methodology is discussed in Part 3. The performance evaluation are described in Part 4, respectively. We present and discuss our results in Part 5, and we draw conclusions in Part 6

Figure 1: Real-time training data set

2. Data description

We use real-time photos as data sets, and we can choose the number of photos we take in real time. This experiment is based on 100 pieces, and then the photos are taken locally,and model training is carried out by using the keras deep learning library.

3. Methodology

Our goal can be achieved with four steps.To begin with, we collect enough face images with laptop camera. Then the labels will be assigned to these images and the label numbers are determined by how many person needed to be classified. Further more, after the the preparation of the data sets, a model will be build using CNN for the future recognize work. Finally, test the model accuracy and apply the trained model in the real-world situation.

3.1. Preparation of the data sets

Using the cross validation principle, the data sets have been divide into three categories. The training data,Validation data and the test data.(train:val=0.7:0.3) Changing the dimension sequence of the image data considering different backen type of the Keras. Turn the label data into one-hot type. Normalize the data image to reduce the inaccuracy brought by different feature value unit usage.

3.2. build the model

The entire CNN model is composed of 18 layers(including 4 convolution layers,5 activation function layers ,2 pooling layers ,3 drop out layers,2 fully connected layers, 1 flatten layer and 1 classification layer). The Masking Convolution have been applied, with 3X3 convolution kernel and the pooling kernel is 2X2. In the activation layer the relu function have been adapted to increase the model nonlinearity and reduce the calculation complexity. The Optimizer is classic SGDStochasitic Gradient Descent)(momentum = 0.6)algorithm and the MSE(mean square error) have been used as loss function. some times the model will overfitt.ing if the model have too many parameters but the training data is limited,such problem can be alleviated by randomly disconnect some neurons, this process is called drop out and is a valid way to prevent overfitting.In this case,the model will randomly drop out 10 percent data during the training

3.3. train the model

The model have been trained for 10 epoch and the appropriate batch size is 48.Such parameter need to be adjusted according to the experiment results,either a too big or too small batch size will influence the training efficiency.During the irateration the learning rate is 1e-6 , the suitable learnig rate is also crucial since it will determine the efficiency of irateration. The loss function will not

converge under a large learning rate and will be too slow to reach the local minimum with a small learning rate.

4. Performance Evaluation

After the training and evaluation of the CNN model, it is time to test the model in the real-time situation.

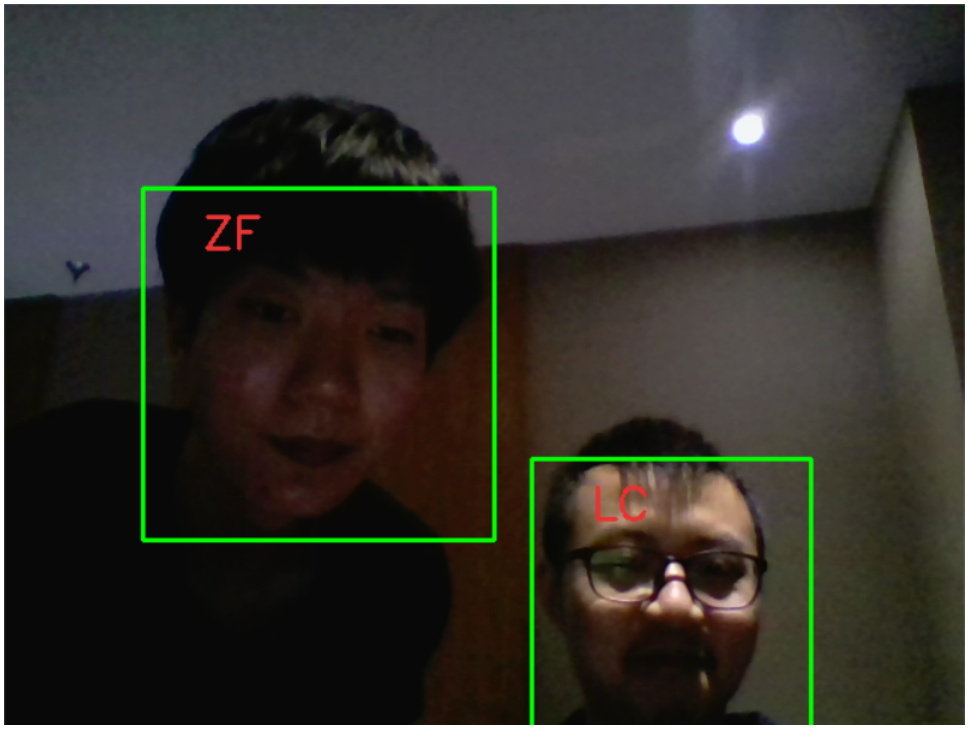

Figure 2: Real-time Performance

5. Results of experiments

After the early stage of the built environment, configuration and obtain real-time video streaming, by calling the opencv identify face, face the data for the training of the model, the convolution neural network is studied and discussed, and finally use keras library in face recognition training model, complete the

final function of face detection and recognition.

Experimental results:

We were conducted for 10 rounds training (nb-epoch = 10), 42 each round iteration (840/20, the training set 1200 x (1-0.3) = 840), each iteration training using 20 samples (batch-size = 20), get the training result is good for 10 rounds training results (for example) :

Training error (loss) : 0.0929.

Accuracy of training (acc) : 0.9093.

Validation error (val-loass) : 0.1177.

Verification accuracy (val-acc) : 0.9517.

The above data is the result of a training. After many experiments, the recognition rate of human face is more than 90 percent.

We then use the test set test model provided by the previous Dataset class.

Finally, after many tests, the accuracy rate is basically over 90 percent, which can accurately distinguish two faces and display the previously typed names.

6. Conclusion

Through this experiment, we can use the opencv, keras libraries to build the different neural networks proficiently, although this project is based on others work, the previous projects have been improved in 3 aspects: 1.many bugs have been fixed, running process have been optimized. 2.increased number of faces to be recongnized, (the old project can only detect 1 person now it can recognize 2 persons) 3. More user friendly, we can type in the person's name more flexiable rather than present captured faces with a fixed name.

But there is small potential bug, all the documents containing the photos will be created before it have been assigned labels. since the labels are assigned to the document from the top one to the bottom one, the faces photos will get the wrong label if the order of the documents is based on the its name rather than its creation time. For instance, document B is first to be created and docu-

ment A is followed, every photos in document B should be assigned label 0,and every photos in the document A should be assigned label 1, but if we oder the document accoding to its name rather than its creation time, the document A which created later will rank first, and all the picutures in document A will be assigned the label 0, this can excange the name of the person to be detected in the future. The solution to this bug is to sort the photo documents according to its creation time rather than its name.

In a word, our group members have benefited a lot from this project and have a clear idea about how to conduct research in the field of deep learning.

Thanks for all the help and guidance provided by Doctor Zhou and Professor Qu!

References

[1] Chellappa, Rama, Charles L Wilson, and Saad Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of the IEEE*, 1995, *83* (5), 705–741.

[2] Lawrence, Steve, C Lee Giles, Ah Chung Tsoi, and Andrew D Back, "Face recognition: A convolutional neural-network approach," *IEEE transactions on neural networks*, 1997, *8* (1), 98–113.

[3] Rowley, Henry A, Shumeet Baluja, and Takeo Kanade, "Human face detection in visual scenes," in "Advances in Neural Information Processing Systems" 1996, pp. 875–881.

[4] Sung, Kah-Kay and Tomaso Poggio, "Learning human face detection in cluttered scenes," in "International Conference on Computer Analysis of Images and Patterns" Springer 1995, pp. 432–439.

[5] Turk, Matthew and Alex Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, 1991, *3* (1), 71–86.