

A Lifelong Learning Approach for Improving Accurate Face Recognition

Zhangqu Yu

School of Software, Shanghai Jiao Tong University, China

Abstract—With the popularity of artificial intelligence and computer vision, an increasing number of software engineers attempt to make their systems be able to recognize the user, through the way of face recognition, e.g., Characteristic Points (i.e., CP)-based face recognition. Generally, the traditional CP-based face recognition only work when the user's face is not much different from the one that stores in the system. However, users' faces can change a lot intentionally or unintentionally, which brings a great challenge for correct face recognition. In view of this challenge, a novel face recognition approach LL-FR (Lifelong Learning-based Face Recognition) is put forward in this paper. Concretely, the system stores not only the face image that the user registered, but also the images every time when it recognizes the user's face. Afterwards, the system makes the future recognition of the user based on all the previous faces of him/her. Finally, through a set of experiments, we demonstrate the feasibility of our approach.

Keywords—face recognition; machine learning; neural network; lifelong learning; artificial intelligence

I. INTRODUCTION

Recently, with the adoption of artificial intelligence and computer vision in both academic and industrial areas, more and more applications now have the requirement to acquire, process, analyze, and understand face images to tell whom this image belongs to. This is called face recognition.

To recognize human faces, many efforts have been made to develop various approaches, e.g., Characteristic Points (i.e., CP)-based face recognition [1]. Concretely, the characteristics of nose, mouth, eyes and et al. are extracted from the face image and compared between the known face and the examining one.

However, the above CP-based face recognition approaches cannot perform very well, when the user don't present the same facial expression with the image that provided to the system. Even if the user make the same pose before the camera, the system still cannot easily recognize him/her when he/she wear glasses, get mustaches, or have longer/shorter hair. In view of this challenge, a novel face recognition approach LL-FR (Lifelong Learning-based Face Recognition) is put forward in this paper. Instead of storing the face image that the user registered the first time, the system stores the images every time when it recognizes the user's face. Afterwards, the system makes the future recognition of the user based on all the previous faces of him/her by machine learning.

The remainder of this paper is organized as follows. In Section 2, we describe the related works and their deficiencies. In Section 3, a novel face recognition approach LL-FR is

brought forth. Evaluation is presented in Section 4. And finally, in Section 5, we summarize the paper and get conclusions.

II. RELATED WORK

A. Characteristic Points-based Face Recognition

This approach analyzes the characteristics of a person's face images input through a digital video camera. It measures the overall facial structure, including distances between eyes, nose, mouth, and jaw edges. These measurements are used as a comparison when a user stands before the camera.

Every face has numerous, distinguishable landmarks, the different peaks and valleys that make up facial features. Each human face has approximately 80 characteristic points. Some of these measured by the Facial Recognition Technology are shown in Figure 1.

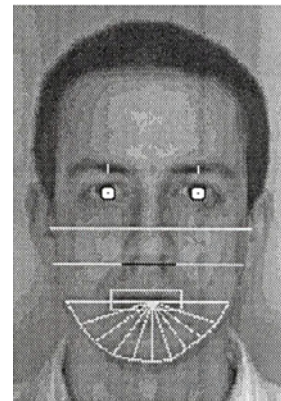


FIGURE 1. GEOMETRICAL FEATURES (WHITE) USED IN THE FACE RECOGNITION EXPERIMENTS

The above characteristic points-based face recognition approaches cannot perform very well, when the user don't present the same facial expression with the image that provided to the system. Even if the user make the same pose before the camera, the system still cannot easily recognize him/her when he/she wear glasses, get mustaches, or have longer/shorter hair.

B. Principal Components Analysis-based Face Recognition

Sirovich and Kirby [2] were the first to utilize Principal Components Analysis (PCA) [3, 4] to economically represent face images. They demonstrated that any particular face can be efficiently represented along the eigenpictures coordinate space (demonstrated in Figure 2), and that any face can be approximately reconstructed by using just a small collection of

eigenpictures and the corresponding projections (‘coefficients’) along each eigenpicture.



FIGURE II. PRINCIPAL COMPONENTS ON SOME FACE DATASET

Turk and Pentland [5, 6] realized, based on Sirovich and Kirby’s findings, that projections along eigenpictures could be used as classification features to recognize faces. They employed this reasoning to develop a face recognition system that builds eigenfaces, which correspond to the eigenvectors associated with the dominant eigenvalues of the known face (patterns) covariance matrix, and then recognizes particular faces by comparing their projections along the eigenfaces to those of the face images of the known individuals. The eigenfaces define a feature space that drastically reduces the dimensionality of the original space, and face identification is carried out in this reduced space.

C. Neural Network-based Face Recognition

A neural network for face recognition is defined by a set of input neurons which may be activated by the pixels of an input image. After being weighted and transformed by a function (determined by the network’s designer), the activations of these neurons are then passed on to other neurons. This process is repeated until finally, an output neuron is activated. This determines whom the input face image belongs to. The neural network model is shown in Figure 3.

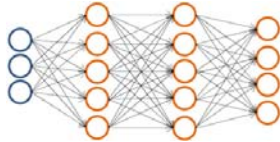


FIGURE III. A TYPICAL ARCHITECTURE OF NEURAL NETWORK

Weng, Ahuja and Huang [7] made use of a hierarchical neural network which was grown automatically and not trained on the traditional gradient descent method. They reported good results on a database of 10 subjects.

Lawrence, Giles, Tsoi and Back [8] reported a 96.2% recognition rate on the ORL database (a database of 400 images of 40 individuals) using a hybrid neural network solution which combines local image sampling, a self-organizing map [9, 10] neural network (which provides dimensionality reduction and invariance to small changes in the image sample), and a convolutional neural network (which provides partial invariance to translation, rotation, scale and deformation).

III. SYSTEM DESIGN

In this section, we introduce a novel approach LL-FR, to recognize face image even if the user present different facial expression, wear accessories, or get old. Concretely, LL-FR consists of four steps in Figure 4.

- Step 1: Get training set and initialize the neural network.↵
- Step 2: Capture face images and recognize the user.↵
- Step 3: Update weights of data and reconstruct training set.↵
- Step 4: Update the neural network.↵

FIGURE IV. FOUR STEPS OF FACE RECOGNITION APPROACH LL-FR

A. Get Training Set and Initialize the Neural Network

In this step, we first take n photos of each $user_i$ ($user_i \in U$). These photos are different from each other in facial expressions, head angles, and other factorial dimensions. Then we tag each photo p_{ik} ($k = 1, 2, \dots, n$) with its user y_i ($y_i \in Y$). Finally, we get $n * ||U||$ tagged photos.

Before we can employ these photos, some preprocessing procedures are needed. For every photo p_{ik} ($i = 1, 2, \dots, ||U||$ and $k = 1, 2, \dots, n$), we detect the location and size of the face and extract this part f_{ik} from the whole photo p_{ik} . Then we transform the colored face photo f_{ik} to gray-scale photo x_{ik} with fixed width W and fixed height H . After preprocessing the origin photos, we get $n * ||U||$ gray-scale photos consisted of $W * H$ pixels and the gray-scale of each pixel is range from 0 to 255.

The neural network of this system consists of 4 layers. The first layer is the input layer which has $W * H$ units. The second and third layer is hidden layers with s_2 and s_3 units respectively. The forth layer is the output layer which has $||U||$ units.

We input X ($x_{ik} \in X$) to this neural network and train the neural network with back propagation algorithm according to the tag we have put on each x_{ik} . After the training is complete, we can get the weight matrix Θ which controls the function mapping from one layer to the next.

B. Capture Face Images and Recognize the User

After the initialization, a camera automatically captures the photo of user. The same preprocess procedures as the above are applied to this photo and the test data $x_{unknown}$ is produced. Using (1) and (2), we can calculate the value y of output layer and find out who this photo belongs to according to the relationship between X and the tags Y .

$$\alpha^{(i)} = g(\Theta^{(i-1)} \alpha^{(i-1)}). \quad (1)$$

$$g(z) = (1 + e^{-z})^{-1}. \quad (2)$$

$\Theta^{(i-1)}$ in (1) is the matrix of weights controlling function mapping from layer $i-1$ to layer i and $\alpha^{(i)}$ is the output of layer i . Equation (2) is the sigmoid function.

C. Update Weights of Data and Reconstruct Training Set

LL-FR stores all photos that is recognized by this system and assigns an initial confidence level c_0 to each of them. The confidence level c_{ik} of every photo p_{ik} follows (3).

$$c_{ik} = c_0 - e^{(\Delta t - \mu)}. \quad (3)$$

The parameter μ in (3) is the interval of current time and the create time of the photo. The hyper parameter Δt in (3) controls the attenuation of confidence level. The confidence level will decrease more rapidly when get Δt larger.

When reconstructing the training set, we select those photos of which the confidence level is higher than some threshold c_r .

D. Update the Neural Network

After the new training set is constructed, the neural network will be trained again like the initialization step. This step will be carried out every night due to the time cost. The new neural network is more likely to recognize users due to the employment of new photos and the abandonment of old ones.

IV. EVALUATION

In this section, we first describe the experiment design. Afterwards, an analysis of experiment results is presented.

A. Experiment Design

The users in this experiment are 8 lab students. In the beginning, 20 photos of each student are taken to initialize the neural network. A camera is placed at the entrance of the front door so that it can easily take photo of users when they come in or go out.

All photos are transformed to $64 * 64$ gray-scale images. The initial confidence level c_0 is 1 and the threshold c_r is 0.05. The hyper parameter Δt is set to 50.

The success and failure face recognitions among 3 months are all logged. The training set reconstruction and the neural network training are carried out at 3 a.m. every day.

B. Experiment Results

We count the successes and failures every month and the result is shown in Table 1.

TABLE I. FACE RECOGNITION COUNT

Time	Recognition Statistics		
	Photos	Successes	Ratio
1st month	912	855	93.75%
2nd month	906	851	93.93%
3rd month	921	866	94.03%

Users' faces are slightly different from the first day when they took photo for initialization. But as we can see in this table, the successful recognition ratio is 93.75% in the first month which is fairly high. In the second and third month, the successful recognition ratio is increased by 0.18% and 0.28% than the first month, respectively. It validates the feasibility of

LL-FR approach on face recognition and the possibility that LL-FR approach can increase the recognition accuracy.

V. CONCLUSIONS

In this paper, a novel face recognition approach LL-FR is put forward, to increase recognition accuracy when users' faces get changed. Instead of comparing characteristic points or making use of the initial face, LL-FR stores and employs all the images when it recognize the user's face. Through a set of experiments, we validate the feasibility of our proposal.

ACKNOWLEDGMENT

I want to extend my gratitude to all those people who help me all the way along. They are my friends in the TCloud lab who participate in this experiment and provide precious data to me. They are my families who keep encouraging me whenever I get stuck in the work. Without them, it will be an impossible mission to finish my research work.

REFERENCES

- [1] R. Brunelli and T. Poggio, "Face recognition: features versus templates," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.15, pp.1042-1052, 1993.
- [2] L. Sirovich and M. Kirby, "Low-dimensional Procedure for the Characterization of Human Faces," Journal of the Optical Society of America A: Optics, Image Science, and Vision, Vol.4, pp.519-524, 1987.
- [3] A. K. Jain and R. C. Dubes, Algorithms for Clustering Data. New Jersey: Prentice-Hall, 1988.
- [4] K. Fukunaga, Introduction to Statistical Pattern Recognition, second ed. Boston, MA: Academic Press, 1990. (PCA)
- [5] M. Turk and A. Pentland, "Face Recognition Using Eigenfaces," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1991, pp.586-591.
- [6] M. Turk and A. Pentland, "Eigenfaces For Recognition," Journal Of Cognitive Neuroscience, Vol.3, pp.71-86, 1991.
- [7] J. Weng, N. Ahuja, and T. S. Huang, "Learning recognition and segmentation of 3-D objects from 3-D images," in Proceedings of the International Conference on Computer Vision (ICCV 93). Berlin, Germany, 1993, pp.121-128.
- [8] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face Recognition: A Convolutional Neural Network Approach," IEEE Transactions on Neural Networks, Special Issue on Neural Networks and Pattern Recognition, pp.1-24, 1997.
- [9] T. Kohonen, "The self-organizing map," Proceedings of the IEEE, Vol.78, pp.1464-1480, 1990.
- [10] T. Kohonen, Self-organizing maps. Berlin, Germany: Springer-Verlag, 1999.