ISBN 978-952-5726-07-7 (Print), 978-952-5726-08-4 (CD-ROM)
Proceedings of the Second Symposium International Computer Science and Computational Technology(ISCSCT '09)
Huangshan, P. R. China, 26-28,Dec. 2009, pp. 222-226

Gender Recognition with Face Images Based on PARCONE Model

Changqin Huang [1], Wei Pan [2], and Shu Lin [2]
[1] Department of Cognitive Science, Xiamen, Fujian, China
Email: cqinhuang@163.com
[2] Department of Cognitive Science, Xiamen, Fujian, China
Email: { wpan@xmu.edu.cn, linshu@ymail.com}

***Abstract*—In this paper, a new type of neural network model——PARCONE (Partially Connected Neural Evolutionary) was proposed, which can overcome the disadvantage that the previous neural networks can not accept more than thousands of inputs. With this new model, no feature extraction is needed before target identification and all of the pixels of a sample image can be used as the inputs of the neural network. After 300 ~ 600 generations' evolution, the new neural network can reach a good recognition rate. With this new model, a gender recognitionwe experiment was make on 490 face images (245 females and 245 males from Color FERET database), in which include not only frontal faces but also the faces rotated from -40°~40°in the direction of horizontal. The gender recognition rate, rejection rate and error rate of the positive examples respectively achieve 95.14%, 2.16% and 2.7%. The experimental results show that the new neural model has a strong pattern recognition ability and can be applied to many other pattern recognitions whick need a large amount of input information.**

***Index Terms*—neural network, PARCONE, face images, gender recognition rate**

I. INTRODUCTION

Gender classification is an important visual tasks for human beings, such as many social interactions critically depend on the correct gender perception. As visual surveillance and human-computer interaction technologies evolve, computer vision systems for gender classification will play an increasing important role in our lives.

As human faces provide important visual information for gender perception, a large number of researchers have investigated gender classification from human faces. In the early years, Moghaddam investigated nonlinear Support Vector Machines (SVM) for gender classification with low-resolution thumbnail face, and demonstrated the superior performance of SVM to other classifiers[1]. Walawalkar adopted SVMs for gender classification using audio and visual cues [2]. Shakhnarovich developed a real-time face detection and demographic analysis (female/male and asian/noasian) system using Adaboost, which delivers slightly better performance than the SVM on unaligned faces from real-world unconstrained video sequences [3].

Recently, various neural network techniques were employed for gender classification from human faces.[4-6]. Due to the size of neural network input vector with the increase of rapid growth, if we identify the image contain face with each pixel as the neural network input, it will make the neural network structure too complex to calculate the right output, and it will also cause badly real-time, or even non-convergence of network and other issues. Currently, researchers often use the methods of image feature extraction (such as border identification, principal components analysis (PCA), etc.). Then they train neural network with the features which are significantly reduced in the dimension of feature space. Despite this method could have been avoided excessive dimension, but to the specific issues, how to choose the characteristics? How many characteristics should be selected? There is not a unified approach. Using the eigenvector as the input of a neural network may be a simple way, but man-made feature selection would lose some of the key information of objectives, so the capacity of identification of neural network would be reduced.

It may be a contradiction that how to make the neural network to take full advantage of images containing important information, while not in a disaster of dimension. Our group proposed a new neural network model -- Model to solve this problem[8]. In the new model, firstly, every neuron can be connects with each other, but only a certain numbers connection values (for example 20) are not zero. Then we identify each image with every pixel value as a neural network input. During training, we change each neuron's previous 20 connections to other 20 connections by the rule of genetic algorithm (GA). For the sake of GA, we can gain the optimum (or sub optimum) 20 connections of every neuron. In this way, our neural network can automatically select the most important features of the objectives, achieve convergence and gain the capacity of target identification.

II. THE PARCONE MODEL

In earlier years, we evolved fully connected neural net modules (that all neurons can be connected with each other and are not limited by the connection number of a neuron.)[7], arguing that they were the "general case". By starting off with every possible connection (i.e. all N^2 of them if there are N neurons in the module) one could let the evolution decide if a particular connection should not exist, by driving down the value of its weighted connection to zero.

AP-PROC-CS-09CN005

This approach was fine, so long as the applications using the fully connected neural net modules did not require too many neurons *N*. If the sum of the pixel points of an image is more than thousands and all these pixels are inputted into the fully connected neural network, the evolvement time of the net will increase up to an unacceptable degree and the net may not achieve convergence. To deal with this problem, currently, there are two methods:

One of the methods is to reduce the size of the input images to an acceptable degree, such as hundreds of pixels, before inputting these images into the network. However, if the number of pixels falls below the 1000 mark, the quality of the image becomes rather poor and makes recognition between subtly different objects difficult. Another method is feature extraction (such as border identification, PCA, etc.), which has been discussed in the part of introduction.

Hence we decided to modify our old neural net model [8] which was fully connected, and make it partially connected.

The Partially Connected Neural Evolutionary (PARCONE) consists of three layers, as shown in Figure 1. They are:

1) Input layer, which has *I* neurons. Each of the neurons can input a pixel of a training image or a testing image. Each neuron of the input layer can be connected to *K* neurons of the Middle layer.

2) Middle layer, which contains *M* neurons. Each neuron of the Middle layer can be connected to *K* neurons of the Input layer, the Middle layer or the Output layer.

3) Output layer, which consists of *O* neurons, and each of the neurons can be connected to *K* Middle layer neurons.

I, *M*, *O* and *K* all can be change depended on the users, and $N = I+M+O$, where *N* is the sum of neurons of the whole net.

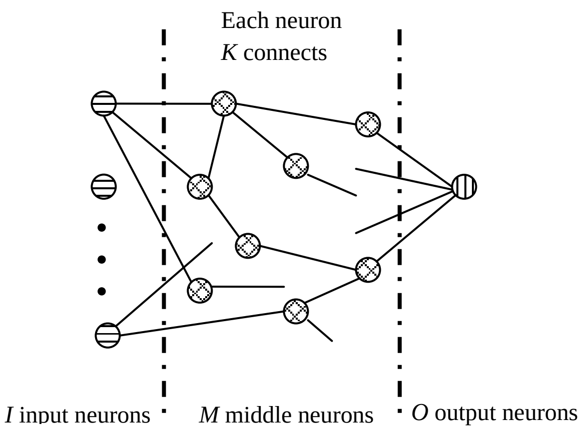

Figure 1. The structure of the new neuron network

Because there are so many neurons in our new network that the traditional standardization index-shaped function, such as:

$$S_i = (2.0/(1.0 + e^{-A_i})) - 1.0 \tag{1}$$

will be invalid. Our experiments show that function (2) can meet the needs of evolutional calculation of our new net [8].

$$S_i = A_i / (|A_i| + c) \tag{2}$$

where $A_i = \sum_{j=1}^{N} W_{ji} S_j$ is an usual active function of neural networks, *c* is a constant.

In PARCONE, for each neuron in the module, to list all the other neurons that that individual neuron connects to. Hence a partially connected neural net model will consist of a list of lists of inner-neural connections, one list per neuron. Each neuron in a module is given a unique non-negative integer ID.

That which specific neurons a neuron connects to and what its connection weight is are continuously adjusted according to the evolutionary algorithm. Using several entire original images (without any feature extraction) where include the objects to be identified as training samples, after 300~600 generations' evolutionary computations, each neuron will eventually be stably connected to *K* other neurons and *K* connecting weights occur, while the whole network acquires the ability to identify a specific object.

Figure 2 shows the data structures used in the coding of the Parcone model. A pointer points to a population of genetic algorithm chromosomes, i.e. pointers to a population of (partially connected) neural network modules. Each pointer in turn points to a further table of pointers, where each pointer points to a hash table for the neural net module in question. The hash table contains pointers to the structs with the ID, weight bits, and weight value. Thus the coding deals with pointers to a nested depth of 4, e.g.

To calculate the output signal of each neuron (at a given moment, called a "tick", where a "tick" is defined to be the time taken for all neurons in the module to calculate their neural output signals) in the module, its hash table is used. A scan across the length of the hash table is performed. From the previous "tick", a signal table of size N (the number of neurons in the module) is filled with the output signals of all the neurons. Assume the first non NULL pointer in the hash table is found at position (slot) "4". The non zero pointer there is used to find the "from" neuron ID. Let us say it is 45. The corresponding weight value might be 0.3452.

At initialization of the genetic algorithm, random connections between the all input neurons and the middle neurons are made, similarly between the middle neurons and other middle neurons or output neuron(s). Each input neuron is connected to a user specified number of middle neurons. Each output neuron is connected to a user specified number of middle neurons [8].

Evolution occurs by mutating the weight values, and/or by creating and deleting random connections.

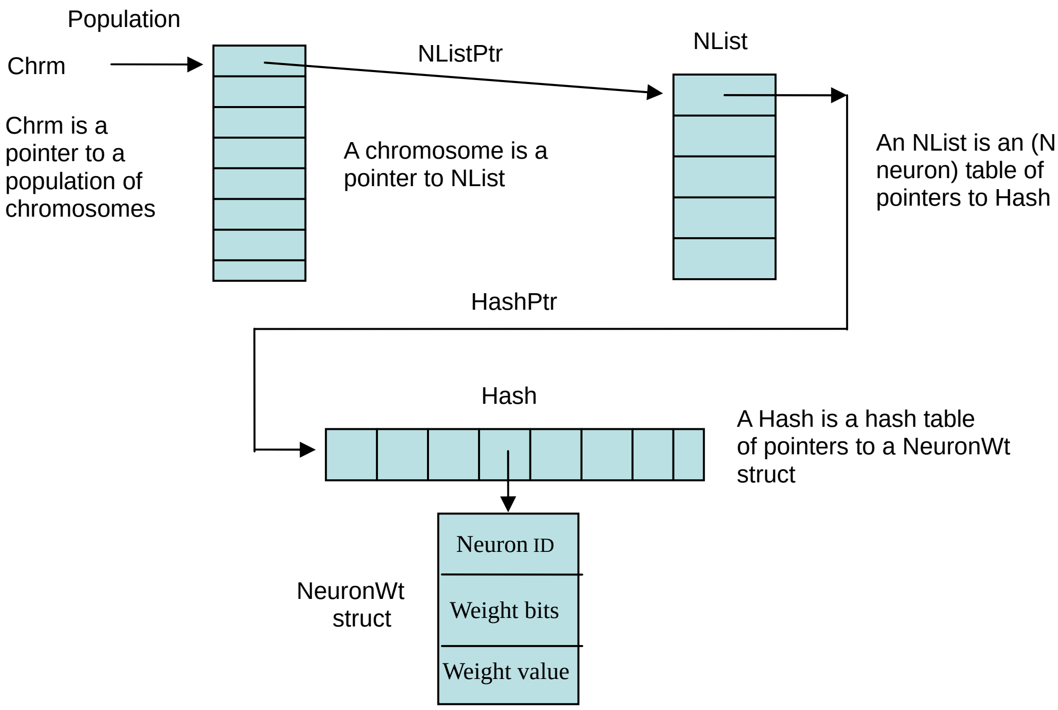

Figure 2. Data Structures of the Parcone Model

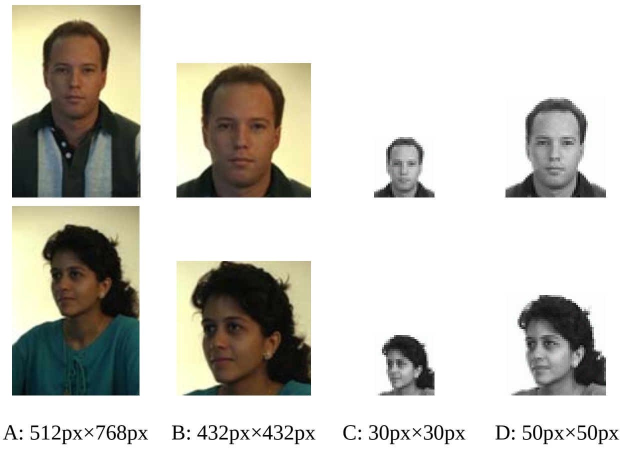

Figure 3. The process of image pre-processing

III. EXPERIMENTAL RESULTS

Our experimental condition includes an ordinary PC machine (CPU: 2.40 GHz; Memory: 1 GB) and software Microsoft Visual C + + 6.0 and OpenCV 1.0.

A. Image Pre-processing

Our experiment uses total 490 face images (245 females and 245 males), randomly withdrawn from the Color FERET database. These images includes not only frontal faces but also the faces rotated from -40°~40°in the direction of horizontal(Figure 3A). Firstly, we cut out the raw images A (Figure 3A) and get images B (Figure 3B). Then, we convert B into gray images. After properly exposure adjusting and special size transformation of these gray images, we get images C (Figure 3C) and D (Figure 3D).

B. Experimental Results

In this paper, three groups of experiments are designed. The *Op* (output value of the neuron network) of each output neuron of our network is a decimal between -1 and 1.

To designate a male test sample as input of the network, when the output of the network Op> 0.1, we set that the identification of the network is correct; when $-0.1<Op<0.1$, we set that the network can not identify the input; when $Op<-0.1$, we set that the identification of the network is wrong.

To designate a female test sample as input of the network, when the output of the network $Op<-0.1$, we set that the identification of the network is correct; when $-0.1<Op<0.1$, we set that the network can not identify the input; when Op> 0.1, we set that the identification of the network is wrong.

Experiment 1: This experiment takes C-type images (30px×30px, Figure 3C) as training samples and testing samples. The training set includes 60 male images as positive examples and 60 female images as counter-examples, while the test set includes 185 male images and 185 female images.

The neural network (Figure 1) consists of three layers, the input layer, the middle layer and the output layer, in which there are I=900(neurons), M=600 (neurons) and O=1 (neuron)respectively. The number of connections of each neuron in the network is K(=20). The experimental results are as shown in table 1.

In table 1, P and N respectively denote positive example and counter-example.

TABLE 1.
THE EXPERIMENTAL RESULTS OF 30 × 30 SAMPLES (%)

Training generations		Recognition rate	Rejection rate	Error rate
300	P	81.62	4.86	13.52
	N	80.54	2.16	17.3
400	P	83.24	10.27	6.49
	N	77.3	3.24	19.46
500	P	69.19	7.03	23.78
	N	77.84	3.24	18.92
600	P	68.65	7.03	24.32
	N	79.46	2.16	18.38

Experiment 2: This experiment uses D-type images (50px ×50px, Figure 3D) as training samples and testing samples. The training set includes 60 male images as positive examples and 60 female images as counter-examples, while the test set includes 185 male images and 185 female images.

The neural network is as shown in Figure 1. The neuron number of the input layer, the middle layer and the output layer is 900, 600 and 1 respectively. The experimental results are as shown in table 2.

TABLE Ⅱ.
THE EXPERIMENTAL RESULTS OF 50 ×50 SAMPLES (%)

Training generations		Recognition rate	Rejection rate	Error rate
300	P	92.43	2.71	4.86
	N	72.97	3.78	23.25
400	P	95.14	2.16	2.7
	N	71.35	1.1	27.55
500	P	96.2	1.1	2.7
	N	71.35	2.16	26.49
600	P	94.06	2.16	3.78
	N	73.51	3.24	23.25

Experiment 3: In this experiment, the training set includes the training set of experiment 2 and the test samples which have not been correctly recognised in experiment 2, while the test set is the same as the test set of experiment 2.

TABLE Ⅲ.
THE EXPERIMENTAL RESULTS OF THE EXPERIMENT 3 (%)

Training generations		Recognition rate	Rejection rate	Error rate
300	P	88.11	3.24	8.65
	N	88.11	5.41	6.48
400	P	96.22	1.08	2.7
	N	92.43	1.62	5.95
500	P	92.43	1.62	5.95
	N	93.51	1.08	5.41
600	P	90.81	2.16	7.03
	N	93.51	2.16	3.23

The neural network is as shown in Figure 1. The neuron number of the input layer, the middle layer and the output layer is 900, 600 and 1 respectively. The experimental results are as shown in table 3.

C. *Analysis of the experimental results*

By the analysis of the experimental results, we get that:

1) Although we use all pixels of a sample image as the input (I = 2500) of the neural network, the novel Partially Connected Neural Evolutionary Model can still achieve convergence. It shows that this network can handle a large number of inputs.

2) The process of the image pre-processing done before the experiment is very simple. In the process, there is no feature extraction and almost all of the information of the original images has been retained. But in this way, not only the target information in the sample images, but also some other interference information, such as glasses, jewelry and different hairstyles, retains. That is why the recognition rates of the counter examples (female) are lower than the recognition rates of the positive examples (male) in experiment 1 and experiment 2.

3) That the face images we adopt for our experiments are not only frontal faces but also the faces rotated from -40°~ 40°in the direction of horizontal is different from the approaches shown in the previous literatures. On the one hand, it increases the difficulty of target identification. On the other hand, it makes our experiments more practical.

4) After 400 generations' evolution, the experimental results can preliminarily meet our expectation. For the sample images of 30px × 30px, the recognition rate of the positive examples is 83.24%, the rejection rate is 10.27% and the error rate is 6.49%, while for the sample images of 50px × 50px, the recognition rate, the rejection rate and the error rate of the positive examples are respectively 95.14%, 2.16% and 2.7%. The above experimental results show that the greater a sample image is, the more information it contains, and the more accurately our network identifies the targets. With the size increase of the sample images, our network can still achieve convergence, but it takes more time for evolution.

5) By adding the previous test samples which can not be identified or can not be identified correctly to previous training set, we get a bigger new training set. To evolve the new training set, the recognition rate can be improved, while the rejection rate and error rate are reduced. As table 3 shows, the recognition rate, the rejection rate and the error rate of the positive examples achieve 96.22%, 1.08% and 2.7% respectively.

6) The disadvantage of our experiment is that the training time is very long. It usually takes 10～30 hours for one experiment. To solve this problem, our group puts forward two possible solutions. One of the solutions is to use computer cluster. Another solution is to use a new introduced tool, Tesla S1070 (960 nuclear calculation module), which has been installed on our LAB. With the

new tool, the evolution speed of our neural network is expected to be increased about 200 times and an experiment can be completed in minutes.

IV CONCLUSIONS

Our objective is gender recognition with face images based on PARCONE model. The PARCONE model is a novel neural network and was first proposed and completed by our group. The results of experiments on human faces rotated from -40°~40°in the direction of horizontal show that our method is more feasible. But a lot of work should be done in the future. For example, how many neurons of each layer of the network is more appropriate, how to effectively improve the evolution speed and how many connections of each neuron in the network will be better.

ACKNOWLEDGMENT

The authors would like to thank Hugo de Garis for theoretic consultation and the Chinese National Natural Science Foundation(60975084) and the Science Foundation of Fujian Province of China(2009J01305) for financial support.

REFERENCES

[1] B.Moghaddam and M.Yang. Learning Gender with Support Faces[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2002,24(5):707–711.

[2] L.Walawalkar, M.Yeasin, A.Narasimhamurthy, R.Sharma. Support vector learning for gender classification using audio and visual cues[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2003, 17(3): 417–439.

[3] G.Shakhnarovich, P.A.Viola, B.Moghaddam. A Unified Learning Framework for Real Time Face Detection and Classification[C]// Proc. IEEE International Conference on Automatic Face and Gesture Recognition,2002:14–21.

[4] Tolba A S. Invariant gender identification[J]. IEEE Trans. Digital Signal Processing, 2001,11(3):222–240.

[5] Rowley H A, Baluja S, Kanade T. Neural Network-Based Face Detection[J]. IEEE Trans Pattern Analysis and Machine Intelligence,1998,201, 20(1) :25–38.

[6] Surendra Ranganath, Krishnamurthy Arun. Face recognition using transform features and neural networks [J]. PatternRecognition,1997,3010,30(10):1615– 1622.

[7] Hugo de GARIS. A "PARtially COnnected Neural Evolutionary" Model Serving as the Basis for Building China's First Artificial Brain[C]// Proceedings of 2008 3rd International Conference on Intelligent System and Knowledge Engineering. China: Xiamen, 2008:9–12.

[8] Hugo de Garis,Michael Korkin. The CAM-Brain Machine (CBM) An FPGA Based Hardware Tool which Evolves a 1000 Neuron Net Circuit Module in Seconds and Updates a 75 Million Neuron Artificial Brain f or Real Time Robot Control[J]. *Neurocomputing*, 2002, 42:35–68.