A new method of face recognition based on integrating the results of different artificial neural networks

MortazaZolfpour- Arokhlo, MohammadnabiOmidvar, MarziehOmidvar, Mohsen Moradi

Abstract— Face recognition is one of the most hot and challengeable technologies, which is based on biometrics, and also one of the most potential technologies[13]. As the most natural and friendly identification method, automatic face recognition has become the important part of the next generation computing technology[15]. This paper present a new method of face recognition based on integrating the results of three different neural networks. This method is not relying on the positions of eyes and lip and even if the face is partially covered, the method appears fault tolerant. We learned that by the help of other face specifications, it could be recognized by an acceptable percentage. All the experiments of the study were carried out based on the ORL (Olivetti Research Laboratory) database. For the selected numbers of 20, 30, and 40 subjects, we came to the results of 87%, 85%, and 83.25% respectively and with time delay of 0.0886 sec per image.

Index Terms—Face recognition, neural network, back propagation algorithm.

I. INTRODUCTION

methods mainly fall into feature based and image based/holistic categories, feature based techniques have emphasis on measuring e.g. distance between eyes, Eye socket depth, cheekbones, width of nose and chin etc, based upon these measured features it makes face representations[11]. Whereas in image-based techniques global image representations are taken into account for recognition purposes e.g. Principal Component Analysis (PCA) and Neural Network technique constitutes image based recognition. The present paper shows that unlike most of common approaches which are mainly concentrated on eyes and lip for face recognition [2,3,4,5,7], the face may be recognized by the help of other specifications even when these organs are covered. For this purpose, we will simulate the human brain’s ability by the help of back propagation algorithm and the integrated results of the experiments in three pioneering neural networks. To design the system, initially some forms including different images and questions were prepared and were applied on a survey population of 200 persons. The questions were around the way to recognize the face of different people and imitating it in the designed system. According to the obtained results, we learned that the role of eyes and lip is not so significant in face recognition and that by the help of other face specifications it could be recognized by an acceptable percentage. Based on the results from the conducted experiments, people seem to use various neural systems in face recognition and come to a fuzzy result from the outcomes of their own neural systems. The present system works exactly like that. The next section is going to discuss the pre-processes on the images and the third section is going to explain the method of image preparation and the structure of the used neural networks. And the fourth section will integrate the results from the three neural networks and eventually the last section treats the conclusion.

II. REQUIRED PRE-PROCESSES

In order to prepare the images to be exposed to the neural networks some pre-processes are required. For all the three neural networks initially the images are turned into the Gray-scale forms and then the image size of 190*160 pixels – which based on trial and error method gives the best results – is applied. Also by applying a 3*3 matrix filter, the image noises are removed. To submit the image to the second neural network, since we are looking from the edges of the image, it should turn into binary format. For the third neural network, in order to lower the computation load, the image size has to shrink into 30*18 pixels.

III. IMAGE PREPARATION AND NEURAL NETWORKS

BP Neural Network is a multi-layered feed forward neural network, which is characterized by its good adaptability[9]. A simple three-layered BP network is composed of input layer, hidden layer and output layer.
An example of a three-layer BP network structure is shown in Figure 1 [10].

In all the neural networks applied, there were the three layers of input, hidden, and output and in all of them the number of outputs was equal to the number of subjects. Also in the first layer the Log sigmoid function and in the second layer the linear function was applied. The training function, which is incorporated in our work, is Gradient descent with momentum (adaptive learning rate) back propagation.

The 'symmetric-points network

The first subject which we focused on for various studies concerning identity recognition was the distribution of recognition specifications over all the face. For identity recognition we decided that unlike the common approaches not to enter all the pixels of the image as the input of the network neurons and act in a way that the created matrix -from the means of image sections- would be the best choice for the first type of neural network [6]. The input of the first network, which we call the symmetric-points network is a 40-subject vector resulted as the mean of the pixels in 40 equal sections of the image which was resulted by dividing the image by 8*5 format and mapping it as a matrix into a 40-subject vector. Dividing the image into 40 sections, which resulted the best outcomes, and also the choice of 8*5 formats were based on trial and error. In this network the input layer included 40 nodes. The hidden layer comprised of 30 nodes which based on the experiments handed in the best outcomes. The network resulted in a high rate of 92% of correct answers. Accordingly, the accuracy coefficient of the so-called symmetric-points network was registered 92 percent. Sample images of one person from the ORL database are shown in Figure 2 (size of shown images are lower than standard for proposed system).

The binary network

The next subject which received extensive study concerning the identity recognition was the problem of an individual’s facial curve. According to the experiences from the primary experiments, we found that people pay a specific attention to the length and width of the face. So for identity recognition we decided that by finding the edges of the face image and entering the data into the network, determine the identity in the output. For this purpose, the given image was made binary to be the input of the network. For this very reason we called the network the ‘binary network’.

1) The network inputs

a) The number of white pixels of the image \( n \)

b) Total distance of all white pixels from the upper edge and the right side, \( P_1 = \sum_{i=1}^{n} a_i \)

c) Total distance of all white pixels from the upper edge and the left side, \( P_2 = \sum_{i=1}^{n} b_i \)

d) Total distance of all white pixels from the lower edge and the left side, \( P_3 = \sum_{i=1}^{n} c_i \)

e) Total distance of all white pixels from the lower edge and the right side, \( P_4 = \sum_{i=1}^{n} d_i \)

f) Total distance of all white pixels from the center of image, \( P_5 = \sum_{i=1}^{n} m_i \)

These 6 elements – P1, P2, P3, P4, P5 and \( n \)- act as the inputs of the network neurons. The hidden layer comprised of 25 nodes which based on the experiments handed in the best outcomes. Figure 3 illustrates the way these elements are computed. The binary network was tested for identity recognition. The accuracy coefficient of the network was equal to 58 percent.

Figure 1. Example of three-layers BP neural network structure

Figure 2. Sample images of one person

Figure 3. The computation method of the 6 elements as the input of the binary network
Adjacency network

Given the results of the primary experiments, we understood that people pay a specific attention to the ears. So we decided to base the identity recognition on the specifications which are significant to people. We experimentally found that in over 95 percent of cases the ear width of the individual is equal to one sixth of their face width so we divided the image as Figure 4. As the neurons input, should be all pixels which separated in small rectangles. But this makes the network input too big. So we decided to reduce the size of image into 30*18 and then divide the image width into 6 sections.

As a result the network will have an input of 2*30*3 pixels, the hidden layer comprised of 60 nodes which based on the experiments handed in the best outcomes and the number of outputs will depend on the number of subjects introduced to the network. We tested the network for different images and came to the acceptable accuracy coefficient of 68 percent.

Merging the Result From the Three Networks

Maybe this question could be raised that why we develop three different neural networks as explained and then we mean to integrate them? The best reason is that we mean to develop an identity recognition system by which recognition of the identity of a person with his face partially covered would be possible. In order to come to the final result, when an image is entered to the network as the input for a test there will be outputs by the networks that when multiplied by the networks coefficient of accuracy and then making the total sum of them, the stand of the greatest resulted figure would be equivalent to the stand of the given image in the network output.

In the Figure 5, the outline of the system is presented. In this system ‘m’ stands for the entered images, ‘Si’ for the ith network coefficient of accuracy, and ‘F’ for the outputs between 1 to 100 (depending on the input image rate of similarity with the database images). Now the resulted F for any image is multiplied by the related network’s coefficient of accuracy, then the results for each image in all the three networks (O, P, Q) are summed and to produce the final outputs of R1 to Rm. The position of the biggest resulted R will be equal to the position of the given image in the output. Of course, if the biggest resulted R is less than 152 –about 70% of similarity- the image will be marked as unknown.

For example:

\[ F1 \times S1 = O1 \text{ for network1} \]
\[ F1 \times S2 = P1 \text{ for network2} \]
\[ F1 \times S3 = Q1 \text{ for network3} \]
\[ R1 = O1 + P1 + Q1 \]
\[ R2 = O2 + P2 + Q2 \]

\[ \cdots \]
\[ \cdots \]
\[ Rm = O_m + P_m + Q_m \]

The proposed system is shown in Figure 5 (size of shown image is lower than standard for proposed system).

### TABLE I. RECOGNITION RATE

<table>
<thead>
<tr>
<th>No. of Subjects</th>
<th>Percentage of Recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 (200 images)</td>
<td>87</td>
</tr>
<tr>
<td>30 (300 images)</td>
<td>85</td>
</tr>
<tr>
<td>40 (400 images)</td>
<td>83.25</td>
</tr>
</tbody>
</table>

The algorithm is implemented using 2.4 GHZ Pentium 4 core i5 machine with Windows 8.1 and MATLAB R2013a as the Development tool. The ORL database has 40 subjects with 10 images per person. Two sets of images are required, one for the training of the neural networks and another set of images upon which testing is done. In this research technique each subject’s out of 10 images picked 5 images as training examples and 5 images as testing set for the neural networks. After the training session, the trained networks are tested upon the unseen images and also upon which training has been done.
In first case, 20 subjects (200 images) are taken as input, and after multiple iterations of neural networks training, in testing deduced 87% results. Similarly for 30 subjects (300 images) and for 40 subjects (400 images) its giving 85% and 83.25% results respectively as shown in table1.

IV. CONCLUSION
The present paper shows that unlike the popular methods which put the main concentration on the eyes and lip for face recognition2,3,4,5, the face could be recognized by the help of other specifications and we came to this result at a relatively high percentage. We did it by the human recognition experiments and simulated the brain ability by obtaining an integrated result from three different neural networks. We came to 92% of correct recognition for the first network, to 58% for the second network, and to 68% for the third network. When we integrated the results of the three networks we came to acceptable results. For 20, 30 and 40 different images which were based on the ORL database the accuracies of 87%, 85% and 83.25% were respectively resulted.

REFERENCES


