A Novel Bidirectional Neural Network for Face Recognition

Jalil Mazloum, Ali Jalali, Javad Amiryan
Electrical and Computer Engineering Department Shahid Beheshti University; G.C
Tehran, Iran
J_Mazloum@sbu.ac.ir, A_Jalali@sbu.ac.ir, Amiryan.j@robocyrus.ir

***Abstract*— The recognition of face images is a complicated problem. Face images are often suffered from variations in brightness, head rotation, facial emotions and so on. Besides, amazing abilities of human brain in face recognition in the presence of these variations, contribute to design face recognition systems based on procedure of human brain. Surveying the recognition and perceptual system of human, shows that, this system has hierarchical and bidirectional structure. Furthermore, the performance of the system would strongly be improved by applying the information of upper layers of face recognition system in interpreting and processing the input data.**

In this paper, novel bidirectional architecture for face recognition inspired by human face recognition system is presented via applying inversion in artificial neural networks (ANN's). In this approach, stored data in the inverse network is applied in the recognition system iteratively and then the correctness of face recognition model has been consequently improved by 8%. The proposed model is able to produce 12 various facial expressions on the output, from only one input expression of each person, after training with AUT database images.

Keywords: Robust face recognition; face variations; bidirectional neural network; inverting neural network

I. INTRODUCTION

Face recognition in pictures and videos is one of the most popular subjects in image processing, pattern recognition and neural networks researches. It is a complicated and difficult problem; however its critical applications in security and military purposes have made wide motivations to build a robust and automatic face recognition system [5]. When a face recognition system, receives a new face image at its input, assigns a correct tag (like the person's name) to the image according to the database information. Now two significant questions need to be surveyed: first, which features of the image should be extracted which consist necessary information to recognize the face, and second, how to categorize a new face image based on prior images. These steps are depicted in Fig. 1.

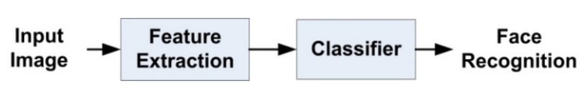

Fig.1. Block diagram of face recognition system

To our knowledge, many approaches have been proposed in pattern recognition fields, for extracting features from face images, and they can be classified into two general groups. First, geometrical methods, which are performing based on features including the nose width, lip border, eyebrow thickness, distance between eyes, and so on [8,9]. These methods aren't reliable as they would perform improperly with any changes in facial expression or environmental circumstances and generally have been used in the past. On the other hand, in recent years, a tendency to use statistical features based on global image appearances has been seen. Also this is supported by new developments in human biology areas. For instance, *principal component analysis* (PCA) and *Fisher linear discriminator analysis* (FLDA) methods, which do a kind of dimension reduction on the data, have a good performance [10-12]. In addition, other techniques including *kernel PCA* (KPCA), *independent component analysis* (ICA), *wavelet transform* (WT), *Zernike and pseudo-Zernike moment invariant*, *neural networks* and etc. have been significantly used for face recognition [15-21].

A lot of researches in order to categorize the face images have been done which use feature vectors. The simplest way is to use Euclidean distance between the extracted feature vector and available feature vectors in the face image database. The minimum distance vector determines that which person is in this image. *Multi-layer-perceptron neural network* (MLP-NN), *radial basic functions* (RBF) neural networks, *support vector machine* (SVM) and other approaches have been utilized for classification [22, 6].

Of course, it goes without saying that, beside these researches, there are still a lot of issues designing a robust and reliable face recognition system. These problems can be considered from two points of view: First, the face images have a lot of similar appearance characteristics; therefore the classification procedure would be difficult. Second, the variations in images including head posture, presence or absence of glasses, hair model and mood of the person and etc. would cause the recognition system to fail. But in most cases, wrong classification is occurred due to face orientation and intensity variations. This matter leads to consider a lot of points during designing a robust face recognition system [7].

We can reduce fault rate of classifier neural networks by deleting these variations from face images, through nonlinear learning property of associative neural networks [24]. Moreover, using *auto-associative* neural network for nonlinear separation of effective variations in face images would result in improvement of the recognition system correctness and virtual images of persons with a desirable variety could be generated as well [25,26].

In this paper, we would introduce a bidirectional model of neural network for recognizing face which is inspired by human face recognition system. The main idea behind this model stems from neurologists' researches; where, they have proved that robust performance of human face recognition system is a consequence of the brain's particular processing procedures [13, 14]. We can mention the ability of mutual information processing as this method's characteristic which is related to the brain.

Bidirectional neural networks, inspired by human brain's processing procedures, are being improved towards performance, flexibility, accuracy and reliability of the human brain. Using these networks in face recognition system helps to estimate and eliminate face variations.

In the following section, Iranian face image database (AUT database) and their preprocessing will be introduced. Section III is dedicated to primary model of face recognition system using ANN. In section IV, the general method for inverting neural networks will be proposed. Section V describes how to improve this bidirectional network; and finally, experimental results will be surveyed.

II. AUT FACE IMAGE DATABASE

Many face image databases have been collected and published by research groups. However the constraints applied in these databases aren't quantitatively and qualitatively identical. For instance parameters including sex, age, race, head rotation and brightness of the image, etc have different distributions in them. A collection of 27 valid face databases in face recognition, face position determination and emotional states, has been mentioned in [1].

We selected and used AUT Iranian face database in this paper. This database has been collected in 2004 by Amirkabir University of technology since there was no appropriate face image database in Iran. But the main reason in collecting AUT image database is that Iranians have particular faces and more important, Iranian women wear mainly veils. This database comprises of 1200 face images of 100 different persons, in which each one has 12 different images with variances in face orientation (looking straight, right and left, up and down), facial expression (smile, anger, wonder, closed-eye), illumination and also presence or absence of glasses. All images have been shot in an identical background with a resolution of 1728×2592 pixels and age variance has been considered as well. An example is shown in Fig .2.

Fig.2. A Sample of AUT database images for one person

We have used the face images of 80 persons (50 men and 30 women) of this database in this paper. The images are employed with the minimum background area (90×120 pixels) as shown in Fig .3.

Fig.3.A sample of face images used for training and evaluation

A. Selecting images for training and test phases

We selected randomly 6 images of 12 available images of each person for training the discussed models in this paper, and used 6 other ones to test them. Thus, 480 images would be used for training and 480 others for test purpose.

B. Preprocessing on the face image data

Each of the AUT face images is a 90×120 matrix which contains intensity of the image pixels. This face image can be assumed as a single vector: $Z_i \subset \Re^{10800\times 1}$. In order to normalize this vector element to range [0, 1], we divided each element of this vector to the maximum gray-level of the images, in the training set, and the following matrix will be derived: $(Z_1, Z_2, \dots, Z_{480}) \subset \Re^{10800\times 480}$

Then the *mean vector* of the face images in training phase can be calculated as follows:

$$\bar{Z} = \frac{1}{480}\sum_{i=1}^{480} Z_i. \qquad (1)$$

Matrix Z, containing all images of training phase, will be calculated from equation 2 and would be used for training the neural network. In this equation, each input face image is normalized into the changes in whole face image database.

$$\boldsymbol{Z} = (Z_1 - \bar{Z}, Z_2 - \bar{Z}, \dots, Z_{480} - \bar{Z}). \qquad (2)$$

In test phase of the models, firstly each face image is converted to a 10800 dimensional vector, and then will be normalized to the range [0, 1]. The difference between these vectors and vector $\bar{Z}$ would be used for evaluation of the models.

III. FACE RECOGNITION MODEL

Supervised feed-forward neural network is frequently used for classification in pattern recognition systems [6,24]. The structure of face recognition model, using neural network, is illustrated in Fig .3. This neural network, with two hidden-layers, can produce decision surfaces with any arbitrary complexity, consisted of cross-planes, for nonlinear mapping of input space (face images) to the classes (persons). The number of input neurons is chosen as the number of input face images' pixels and the number of output neurons is as the number of the classes. We chose sigmoid function as the activation function of linear neurons in the hidden and output layers. The number of nodes in hidden layers is set based on the number of learning patterns. Selecting a small numbers of neurons, leads to weaken generalization feature of the network and lower the model credit. On the other hand, selecting too many neurons would enlarge the size of the model and long the training time, while has trifle effect on the accuracy of the model. These numbers are determined by experiment. N1= 300, N2=200.

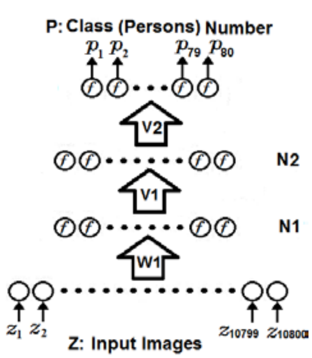

Fig.3. ANN architecture for face recognition model

Corresponding to each face image vector at input(Z_i), a unique code vector $P_i = (p, p_2, \dots, p_{80})^T \subset \Re^{80\times 1}$ is assumed at the output of the network. The aim of learning procedure is to find weights $W_1, V_1 \, and \, V_2$ for mapping: $Z \rightarrow P$ where matrix $\boldsymbol{P} = (P_1, P_2, \dots, P_{480})$ includes all considered codes for input images.

In training phase, we intend to minimize the sparseness within each classes, and maximize it among the classes, therefore the input face image related to a person should set the corresponding output neuron to 1, and other neurons to 0 (according to equation 3).(one and zero are the maximum and minimum values of sigmoid function respectively.)

$$p_i = \begin{cases} 1 & if(Z_i \in C_i) \\ 0 & Otherwise \end{cases} \qquad (3)$$

The learning phase can be done in error back-propagation manner .Thus the error, $\delta_P = f' \times (P_i - \hat{P}_i)$ of the output layer could be used for correction of the network weights, where $\hat{P}_i$ is output vector of the neural network, which is calculated for the input image that belongs to the ith person; And $f' = \hat{P}_i \times (1 - \hat{P}_i)$ is the derivative of the sigmoid function. The learning process will continue until minimizing the energy function according to the equation 4.

$$E_l = \frac{1}{2} \times \sum_{i=1}^{480} \left\| P_i - \hat{P}_i \right\|^2. \qquad (4)$$

In test phase, each face image is fed to the network, and then assigned to the class which its corresponding output node has largest value.

IV. INVERTING THE NEURAL NETWORK OF FACE RECOGNITION MODEL TO FORM THE BIDIRECTIONAL MODEL

The neurologists' researches show that, there are bidirectional connections and interactions between primary and upper layers of sensing processes in brain. It seems that cortex, when receives a sensory signal, add a brief data to it, (in addition of up-to-down processes which is done by neurons) so that, reformed assumptions and imaginations of the person would near to the real objects [13, 14].

Mono-directional model of face recognition system, in section II, makes only a mapping from face images to the person codes, so this network might make a slip in confronting a new case.

Now, we propose a new model of neural network which is inspired of calculative procedures of human brain. In this model, we invert the neural network to obtain the input from the output. The simplest way to implement this inversion is to use two feed-forward networks contrarily [23]. This structure is shown in Fig. 4. The learning process of such a networks would converge to a solution if learning data are one-to-one; however if learning data are many-to-one (like recognition model), the direct network would converge, but the inverse one just reforms the inputs mean.

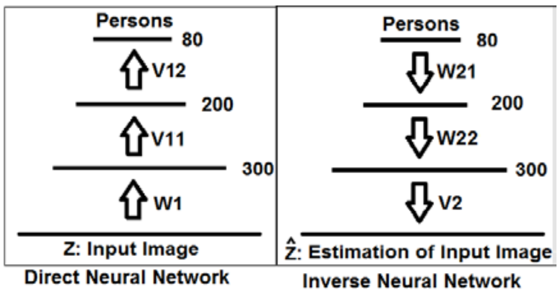

Fig.4. 2 two-layered feed-forward neural networks

Like face recognition model proposed in section III, learning process of the inverse network can be accomplished by error back-propagation. In this network, an estimation of network output can be achieved by putting each person's code on the input. Correction of the network weights will be done by propagating the error $\delta_z = (Z_i - \hat{Z}_i)$ towards the input. The network training process, for 480 images belonging to 80 persons of training set, will continue until the energy function is minimized according to equation 5.

$$E_2 = \frac{1}{2} \times \sum_{i=1}^{480} \left\| Z_i - \hat{Z}_i \right\|^2. \qquad (5)$$

After this learning, 80 virtual images of database persons can be generated by putting code of each person on the input of inverse network. The images generated by inverse network for 6 foremost persons of AUT database are shown in Fig. 5.

Fig.5. Virtual images of 6 foremost persons of AUT database

After completing training phase of two networks, in test phase, each input image that is put on the input of direct network makes a primary recognition at the output; then this output is fed to the input of the inverse network and an estimation of input will be generated at the output of the network. By applying this output to input of the direct network and repeating this, until the network produces the same image, at network input, the ultimate recognition will be done. Figure 6 depicts this procedure.

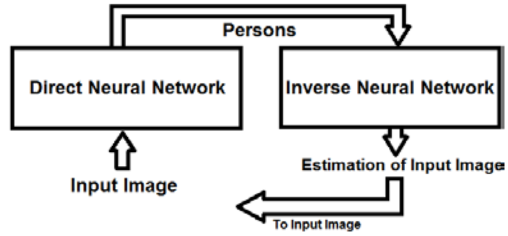

Fig.6. Bidirectional model with reiterating the output

V. Improvement of Proposed Bidirectional Network

The model described in section 4, is the simplest way to implement a bidirectional face recognition model. To improve this model, in addition of each person's information, the information related to the variance in facial expressions should be recognized; for instance when you hear the name of a familiar person, you can remember his/her face with all diversities of facial expression. Further, in the case of unfamiliar persons, the diversities and their types are recognizable. Humans can imagine virtual images of a familiar person with diversities in expressions which have not seen before. (For example adding beard or glasses to the face of a person we have not ever seen)

These images can be imagined, since the brain uses its preceding knowledge like beard or glasses that has seen on face of other persons. Fig.7 presents the structure of a bidirectional neural network based on face recognition system in human.

The direct neural network makes a mapping from face images to the persons and their expression variations. Moreover, the inverse network can approximately reform the images by using persons' information and facial expression diversities.

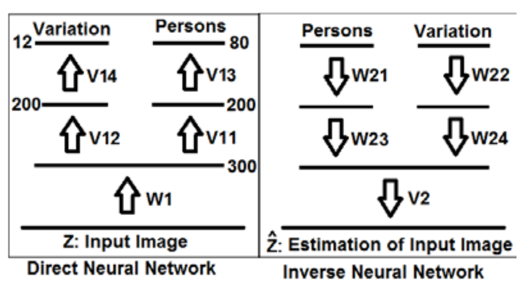

Fig.7. Proposed bidirectional architecture for face recognition

As 12 controlled various facial expressions exist in AUT database, the inverse network is perfectly trained, i.e. after training the inverse network we can regenerate all 480 virtual images of the persons in AUT database by putting the persons' codes on the network input. Furthermore, by the knowledge of different persons of training database, stored in the network, we are able to obtain virtual images of persons while these images don't belong to the database. An example is shown in Fig. 8 for one of the persons of the database. In this image, the inverse network has stored the features, such as smile, beard, and etc of the training database images and puts them on virtual face of this person.

Fig. 8: Virtual images of facial expression diversities for one of persons of AUT database.
These 6 images don't belong to the training set.

As we know the recognition of front facial images without any facial emotion, is easy for human and even computer; so, only normal state images enter the direct network in reiterating the output. Then, like section VI, after training these two networks, each input image would pass through the network and make a primary recognition at the network output. This output will be fed to the input of the inverse network with a normal state and an estimation of the input is generated at the network output. The ultimate recognition will be done by applying this output to the input of the direct network, and repeating this task. This process will continue until the output of inverse network converges. These tasks are shown in Fig. 9 and the results are brought in table 1.

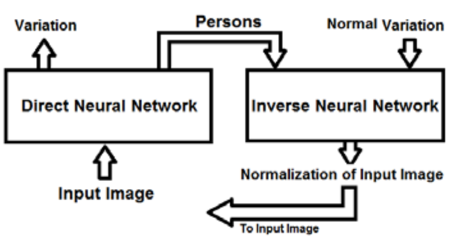

Fig.9. Proposed bidirectional model with state normalization and reiterating the output at input

VI. RESULTS

In this section, the results of the experiments on all models discussed in the paper have been presented. All of the simulations have been done in MATLAB environment running on a dual core Pentium4 with 2GB of RAM. In all proposed networks, the learning rate is set to 0.1, momentum coefficient is 0.7 and all initial weights are chosen 0.1×Randn. To speed up the learning process, while network training, input data are randomly fed to the network and the learning rate reduces with error reduction. In the linear layer of the networks, the function derivation replaced with a small coefficient (0.01), otherwise the learning rate should be set to 0.001 so that the neural network converges. The maximum time for training the proposed bidirectional network is about 12 hours, and the minimum for the recognition model is 50 minutes. The percentages of correctness of the models are shown in table 1.

Table 1.Results of the correctness face recognition models

Face recognition technique	Correctness percent
Recognition Model in Fig. 3 [24]	84.4%
Recognition model in Fig. 4 with reiterating the output at input	86%
direct model Proposed in Fig. 7	84.8%
Proposed bidirectional model with nonlinear normalization and reiterating the output in input	92.5%

VII. CONCLUSION

The usual method of face recognition includes feature extraction from face database and feature classification. Most of the techniques don't use high abilities of human brain in face recognition. Our purpose is to employ the achievements of medical researches and propose models based on human vision system.

Neurologists' investigations about human brain shows that, processes in vision path and the cortex are consisted of feed-forward computations and receiving feedback from higher layers. The recognition task in human isn't performed in one step, but also it is done in several feed-forward and recursive steps and what is recognized in each step, is employed in order to model correspondence in next steps.

Mono-directional architectures of multi-layer feed-forward neural networks, although have the capability of nonlinear mapping, but they aren't efficient in presence of variations in input patterns. Utilizing a feed-forward network beside an inverse network causes to constitute a bidirectional network. This structure increases the abilities of the network in processing and recognition. In this bidirectional structure, stored data in the direct network weights is the knowledge, generated from similarities learned in input patterns, while inverse network learns to reform the same image by available information in decision regions.

Proposed bidirectional architecture, recognizes the persons' information and their variations. This architecture learns the common features in the database. So it can recognize the face images in testing set and even generate virtual images with variations and delete or change these variations.

ACKNOWLEDGMENT

The authors gratefully acknowledge the endeavors of Dr. Karim Faez and specialists at intelligent signal processing lab of Amirkabir University of technology for dedicating AUT database.

REFERENCES

[1] R. Gross, Handbook of Face Recognition, *Springer Verlag*, 2005.
[2] G. Shakhnarovich, B. Moghaddam, "Face Recognition in Subspaces," Handbook of Face Recognition, *Springer Verlag*, 2004.
[3] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks", *Science Mag.*, Vol. 313, 2006.
[4] M. A. Turk and A. P. Pentland, "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, Vol. 3, pp. 71–86, 1991.
[5] R. Chellappa, C. L. Wilson, S. Sirohey, "Human and machine recognition of faces: A survey", *Proceeding of the IEEE*, Vol. 83, pp. 705-740, 1995.
[6] G. P. Zhang, "Neural Network for Classification: A Survey", IEEE Trans. System, Man, and Cybernetics. Application and Reviews, Vol. 30, No. 4, pp. 451-462, 2000.
[7] Y. Adini, Y. Moses, S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 18, pp. 721–732, 1997.
[8] R. Brunelli and T. Poggio, "Face recognition: Features versus templates", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 15, pp. 1042–1053, 1993.
[9] P. J. Phillips, "Matching pursuit filters applied to face identification", *IEEE Trans. Image Processing*, Vol. 7, pp. 1150–1164, 1998.
[10] Z. Hong, "Algebraic feature extraction of image for recognition", *IEEE Trans. Pattern Recognition*, Vol. 24, pp. 211–219, 1991.
[11] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces versus Fisherfaces: Recognition using class specific linear projection", *IEEE Trans. Pattern Analisis and Machine Intelligence*, Vol. 19, pp. 711–720, 1997.
[12] D. L. Swets and J. Weng, "Using Discriminant eigenfeatures for image retrieval", *IEEE Trans. Pattern Analisis and Machine Intelligence*, Vol. 18, pp. 831–836, 1996.
[13] E. Koerner and G. Matsumoto, "Cortical architecture and self-referential control for brain like computation, a new approach to understanding how the brain organizes computation", *IEEE Trans. Engineering in Medicine and Biology Magazine*, 2002.
[14] J. Ghosen and Y. Bengio, "Bias Learning, Knowledge sharing", *IEEE Trans. Neural Networks*, Vol. 14, No. 4. 2003.
[15] B. L. Zhang and Y. Guo, "Face recognition by wavelet domain associative memory", *Intelligenct Multimedia, Video and Speech Processing*, pp. 481–485, 2001.
[16] M. R. M Rizk and A.Taha, "Analysis of Neural Networks for face Recognition systems with feature Extraction to Develop An Eye localization Based Method", *Electronics, Circuits and System*, Vol. 3, pp. 847-850, 2002.
[17] Y. Zhang and C. Liu, "Face Recognition using Kernel Principal Component Analysis and Genetic Algorithms", Neural Network for Signal Processing, pp. 337-343, 2002.
[18] Y. Jian and D. Zhang, "Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 26, no. 1, 2004.
[19] N. kwak, H. Choi, N. Ahuja, "Face Recognition using feature extraction based on Independent Component analysis", *IEEE Trans. Image Processing*, Vol. 2, pp. 337-340, 2002.
[20] J. Haddadnia, K. Faez, M. Ahmadi, "A Fuzzy hybrid learning algorithm for radial basis function neural network with application in human face recognition", *Pattern Recognition*, 2003.
[21] C. Liu and H. Wechsler, "Independent Component Analysis of Gabor Features for Face Recognition", *IEEE Trans. Neural Network*, Vol. 14, No.4, 2003.
[22] J. E. Meng, Wu Shiqian, Lu Juwei, "Face Recognition With RadialBasis Function (RBF) Neural Networks", *Neural Network*, Vol. 3, pp. 2162-2167, 2002.
[23] C. A. Jensen, R. D. Reed, R. J. Marks, "*Inversion of Feedforward Neural Networks: Algorithms and Application*", proceeding of the IEEE Neural Network, Vol. 87, pp. 1536-1549, 1999.
[24] J. Mazloum, S. A. Seyed Salehi, M. Slamifar, "Improve Face Recognition with Variations Removed by the Neural Network", 15^{th} Iranian Conference on Computer Society (CSI 2010 in Persian).
[25] J. Mazloum, A. Jalali, "Robust feature extraction for face recognition using neural networks", 20^{th} Iranian Conference on Electrical Engineering (ICEE 2012 in Persian).
[26] J. Mazloum, S. Khodambashi, A. Jalali, A. Zakerolhosseini, "Development of an Auto-associative Neural Network for Improvement in Face Recognition", 4^{th} IEEE Internatioal Conference on Computer Science and Information Technology (ICSSIT 2011), pp. 343-347, 2011.