MLP Neural Network Based Face Recognition System Using Constructive Training algorithm

Hayet Boughrara[1], Mohamed Chtourou[1], Chokri Ben Amar[2],

[1] Intelligent Control, design and Optimization of complex Systems (ICOS), Sfax Engineering School, BP W, 3038 Sfax,
[2] Research Group on Intelligent Machines (REGIM), University of Sfax, Sfax Engineering School, BP W, 3038 Sfax

hayet.boughrara@laposte.net, mohamed.chtourou@enis.rnu.tn, chokri.benamar@enis.rnu.tn,

Abstract*—Face recognition is one of the most efficient applications of computer authentication and pattern recognition. Therefore it attracts significant attention of researchers. In the past decades, many feature extraction algorithms have been proposed. In this paper Gabor features and Zernike moment were used to extract features from human face images for recognition application. This paper is a study for new constructive training algorithm for Multi Layer Perceptron (MLP) which is applied to face recognition application. An incremental training procedure was employed where the training patterns are learned incrementally. This algorithm started with a single training pattern and a single hidden-layer using one neuron. During neural network training, the hidden neuron is increased when the Mean Square Error (MSE) of the Training Data (TD) is not reduced or the algorithm gets stuck in a local minimum. Input patterns are trained incrementally (one by one) until all patterns of TD are selected and trained. Face recognition system structure based on a MLP neural network was constructed and was tested for face recognition. The proposed approach was tested on the UMIST database. Experimental results indicate that we can obtain an optimal architecture of neural network classifier (with the least possible number of hidden neuron) using our present constructive algorithm, and prove the effectiveness of the proposed method compared to the MLP architecture with back-propagation algorithm.

Keywords-component; face recognition; constructive training algorithm; MLP neural network; Feature extraction; Zernike Moment; Gabor feature.

I. INTRODUCTION

Recently, Biometric identification technologies are more and more used in practical applications. Compared to other biometric technologies, face recognition provides more natural and easier approach human identification. Thus, face recognition is one of the most important parts in biometrics methods identifying individuals by the features of face. There are several successful practical applications of face recognition, such as biometric personal identification, man-machine communication and access control [1].

Many techniques for face recognition are developed. These algorithms can be categorized into two main groups: holistic or global feature and local approaches. In holistic approach, the whole face region is saved as input data into face detection system. Examples of holistic methods are eigenface [2], fisher face [3], LBP faces [4] and so on. In local approaches, local features on face such as nose, and then eyes are segmented and then used as input data for structural classifier. Also, pure geometry, dynamic link architecture, and hidden Markov model methods belong to this category.

As a typical pattern recognition problem, the performance of a face recognition system depends not only on the classifier, but also on the representation of the face patterns. Gabor features have drawn much research attention on the field of face recognition due to its promising biological relevance that the receptive field profiles of neurons in the primary visual cortex of mammals are oriented and have characteristic spatial frequencies. The Gabor feature is robust to variations due to expression and illumination changes and is one of the most successful approaches for face recognition.

The orthogonal moments that are invariant to the basic image transformation such as translation, scaling, and rotation have great importance in pattern recognition. This paper addresses Gabor features and Zernike moment to represent the face image.

The artificial neural networks have been widely used as the classifier in many face recognition systems. The number of hidden neurons layer is one of the most important considerations when solving problems using multilayered feed forward neural networks.

The contribution of this paper is to use the artificial neural network to classify the face using a new proposed constructive algorithm for Multi Layer Perceptron (MLP). In our approach, input patterns are trained incrementally (one by one) until all patterns of Training Data (TD) are selected and trained. Starting with a neural network containing one hidden neuron, we train the network using variants of the back propagation algorithm. The hidden neurons grow during the training when the Mean Square Error (MSE) of the Training Data (TD) is not reduced or the algorithm gets stuck in a local minimum.

In particular, we develop a structure of face recognition system based on feed forward neural networks in order to recognize faces and using the proposed constructive algorithm.

In a first step, we describe the feature extraction techniques using in this study. Then, the constructive training algorithm is illustrated. Next, the face recognition system is describes. Finally, the experimental results achieved on the UMIST database are presented and analyzed.

II. FEATURES EXTRACTION TECHNIQUES

Two central issues to an automatic face recognition system exist: the feature selection for face representation and the classification of a new face image based on the chosen feature representation. The aim of the feature extractor is to produce a feature vector containing all pertinent information about the face while having a low dimensionality. In order to design a good face recognition system, the choice of feature extractor is very crucial.

This paper uses the Zernike moment and the Gabor features to represent the face image and produces recognition task in tandem with neural network.

A. Zernike moments

Statistical-based approaches for feature extraction such as moment invariants [5] have received considerable attention in recent years for their invariance properties. Moment features are invariant under scaling, translation rotation and reflection.

The invariant properties of moments are utilized as pattern sensitive features in classification and recognition applications [6]. Zernike moments are the most widely used family of orthogonal moments due to their extra property of being invariant to an arbitrary rotation of the object that they describe.

The Zernike moments are orthogonal set of complex valued polynomials defined over the polar coordinates inside a unit circle. The Zernike moment of order n and repetition m is defined as [7, 8]:

$$\begin{cases} A_{nm} = \dfrac{n+1}{\pi} \sum_{x} \sum_{y} f(x,y)\, V_{nm}(x,y) \\ \qquad x^2 + y^2 \leq 1 \end{cases} \quad (1)$$

Were $V_{nm}(x,y)$ denote Zernike polynomials of order n and repetition m and is written as:

$$V_{nm}(x,y) = V_{nm}(\rho,\theta) = R_{nm}(\rho) \quad (2)$$

The radial polynomial $R_{nm}(\rho)$ is identified as:

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{n-|m|}{2}} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!} \rho^{n-2s} \quad (3)$$

Were n is a non-negative integer, and m is an integer such that $n-|m|$ is even and $|m| \leq n$, $\rho = \sqrt{x^2+y^2}$ and $\theta = tan^{-1}\frac{y}{x}$

Zernike moment is used as the feature extractor whereby the order is varied to achieve the optimal classification performance.

B. Gabor Feature Extraction

Face representation using Gabor features has attracted considerable attention in computer vision, image processing, pattern recognition, etc [9,10]. Researchers have shown that Gabor features, when appropriately designed, are invariant against translation, rotation and scale [11].

The Gabor features exhibit desirable characteristics of spatial locality, orientation selectively and optimally localized in the space and frequency domains. Generally, the Gabor filters can be defined as follows:

$$\Psi_{\mu,\nu}(x,y) = \frac{k_{\mu,\nu}^2}{\sigma^2} \exp\left(-\frac{k_{\mu,\nu}^2\,(x^2+y^2)}{2\sigma^2}\right). \left[\exp\,(ik_{\mu,\nu}.\begin{pmatrix}x\\y\end{pmatrix}) - \exp\,(-\frac{\sigma^2}{2})\right] \quad (4)$$

Where μ and ν define the orientation and scale of the Gabor filters respectively, and the wave vector $k_{\mu,\nu}$ is defined as follows:

$$k_{\mu,\nu} = k_\nu e^{i\varphi_\mu} \quad (5)$$

Where $k_\nu = k_{max}/f^\nu$, $f = \sqrt{2}$, $\varphi_\mu = \mu\pi/8$, k_{max} is the maximum frequency, and f is the spacing factor between filters in the frequency domain [12].

The Gabor representation of a face image is computed by convolving the face image with the Gabor filters. Let $I(x,y)$ be the intensity at the coordinate (*x*, *y*) in a gray scale face image, its convolution with a Gabor filter is defined as:

$$G_{\mu,\nu}(x,y) = I(x,y) * \psi_{\mu,\nu}(x,y) \quad (6)$$

Where * denotes the convolution operator

The response to each Gabor filter representation is a complex function with a real part and an imaginary part. In this study, the magnitude response is used to represent the features with the following parameters: $k_{max} = \pi$, $f = \sqrt{2}$ and $\sigma = \pi$.

III. CONSTRUCTIVE TRAINING ALGORITHM

The number of hidden layer neurons is one of the most important considerations when solving problems using multilayer feed forward neural networks. The problem resides in the answer to the following question: how much training data do we need to be able to construct a good classifier?

In fact, increasing the size of the database nearly always results in a decrease of error rate and eliminate the over training phenomena.

A constructive algorithm with incremental training has been proposed in order to train the neural network and to take the best architecture face recognition system with a small number of hidden neurons as possible corresponding to the minimum number of generalization error (GE) [13].

- ***Step 1:*** Create the MLP composed by one hidden neuron
- ***Step 2:*** Initialize the weights of layers connections with a random values.
- ***Step3:*** Selecting one input pattern (N_input = 1) from the training data (TD)
- ***Step 4:*** Training MLP with one input pattern selected from the training data (TD) to achieve the system error tolerance specified ε.
- ***Step 5:*** If the training algorithm can reduce the MSE to within Eps, go to Step 6; otherwise, go to Step 7.
- ***Step 6:*** While the N_input < N_tot patterns from TD, increase the number of pattern from the TD by one (N_input = N_input + 1) and go to Step 4; otherwise go to Step 10.
- ***Step 7:*** Store the last weights of layer connections input.
- ***Step 8:*** Increase the number of hidden neurons by one (N_hid = N_hid+1).
- ***Step 9:*** Assign the initial weights and go to Step 3.
- ***Step 10***: Take the best architecture which have a minimum number of generalization error corresponding to a minimum number of neurons.

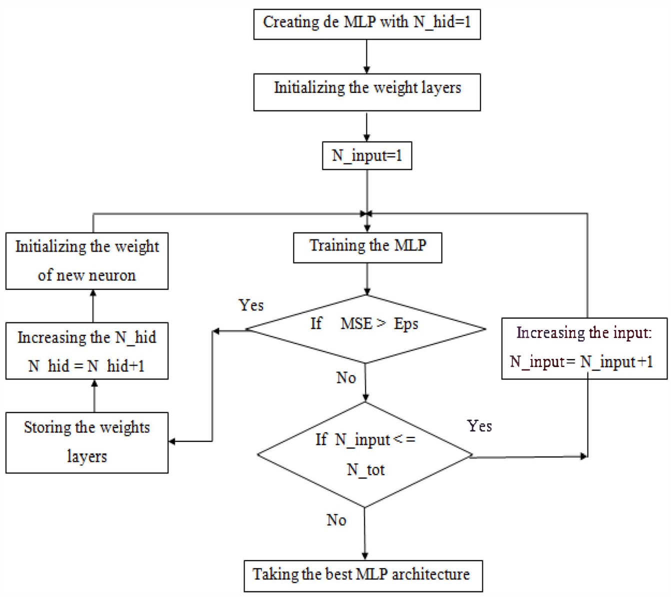

Figure 1. Constructive training algorithm for the MLP

IV. FACE RECOGNITION SYSTEM

Having the feature vectors for all the samples in the train and test sets, the next step would be to design a classifier. In this study, a multi layer Perceptron (MLP) architecture has been used as a classifier of face characterized by the Zernike moment and the GABOR features. The architecture of the used neural network is represented on Figure 2.

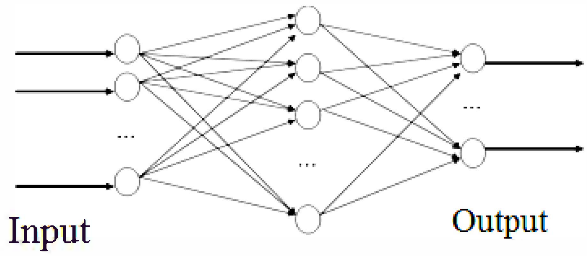

Figure 2. Structure of the MLP networks

We have used a three layer perceptron neural network. The number of neurons in input layer is equal to the size of related feature vector for each experiment and the number of output neurons is equal to the number of individual faces the network is required to recognize [14].

The hidden layer is calculated using a constructive algorithm with the proposed incremental training (see Figure 1.

The MLP networks trained using back-propagation algorithms are used in this work. The inputs vectors are applied to input units that have linear transfer functions. Others units have typically a sigmoid nonlinear function.

The Back-propagation network undergoes a supervised learning process, and the output signal goes through an activation function. The training of MLP network was based on the two following equations:

$$\Delta\omega_{ji}(t) = -\eta \frac{\partial E(t)}{\partial \omega_{ji}(t)} \qquad (7)$$

$$E_p = \frac{1}{2}\sum_k \left(d_{pk} - s_{pk}\right)^2 \qquad (8)$$

Were k is the number of neuron in the MLP outputs and η is the learning rate. E designed the error of the network (Mean Square Error "MSE")

The performance of the recognition system has been measured in terms of recognition rate on testing:

$$R = \frac{N * 100}{Nt} \quad (9)$$

Were N = Number of true classification and Nt = The total number of image in the testing data.

V. EXPERIMENTS

The proposed approach was experimentally evaluated using the UMIST database [14].

A. Database:

The UMIST database consists of 1012 images of 20 people covering a range of poses from profile to frontal views. Subjects cover a range of race, gender, and appearance [15].

UMIST faces is the benchmark database for multi-view face recognition. Each individual is shown in a range of poses from profile to frontal. Figure 3 illustrates images selected for a subject.

The cropped images are in PGM format each with size 220 x 220 pixels with 256-bit grey-scale. For our experiments, we randomly select 633 images as the training set, 40 images as the validation set, and the remaining images as the test set.

In order to enhance the global contrast of the images, and reduce the effect of uneven illuminations, histogram equalization is applied to all the images in our experiment.

Figure 3 - Examples of selected image of the UMIST database.

The feature extraction is based on two methods: the Zernike moments and the GABOR feature. The number of input neuron is equal to the number of the feature extraction. The number of the output layers is equal to the number of class which is 20.

B. Results using the Zernike Moments

In the first time, Zernike moment is used as the feature extractor whereby the order is varied to achieve the optimal classification performance. In our experiment, the best results were obtained for the Zernike moment of order 8. For the classification step, we compared the fix architecture and the constructive training algorithm.

Table 1 gives the recognition accuracies using the Zernike moment of order 8 and the MLP for classification.

TABLE.1. RECOGNITION RATES ON THE UMIST DATABASE USING ZERNIKE MOMENT AND MLP

Number hidden neuron	**Recognition Rate (%)**
10	89,38
15	91,15
20	91,74
25	91,44
30	89,38
35	91,15
40	91,15
45	90,26
50	89,67

TABLE.2. RECOGNITION RATES USING THE ZERNIKE MOMENT AND THE CONSTRUCTIVE TRAINING ALGORITHM.

Eps	**epochs**	**N_hid**	**Recognition rate (%)**
0.01	5000	28	83.48
0.01	10000	45	85.54
0.01	15000	28	81.41
0.01	20000	18	82.59
0.007	5000	15	90.85
0.007	10000	22	94.98
0.007	15000	63	89.50
0.007	20000	18	86.43
0.005	5000	92	87.9
0.005	10000	81	97.34
0.005	15000	28	96.16
0.005	20000	25	91.15

Table 2 shows the recognition rate using the Zernike moment and the constructive training algorithm. We can notice that the result using the constructive training algorithm is higer then the MLP. For the MLP, the rate recognition is 91.74 % with 20 hidden neurons, and for our approaches, the rate is higer.

C. Results using the GABOR feature.

The Gabor representation of an image involves convolution of the image with a family of Gabor filters at different spatial frequencies and different orientations.

Firstly, the Gabor faces, representing one face image, is computed by convoluting it with a Gabor filters with one orientation and one scale. Figure 2 shows an example of the original image and the image using the GABOR filters using only the magnitude value. Then, The Personal component analyzing PCA (eigenfaces) for dimensionality reduction is applied on the convolved image. Thus, we had getting the feature matrix which is the input vector of the MLP classifier.

Figure 4 – The original image and the magnitude of the convolution with a Gabor filters

Table 2 presented the recognition rates according to the Gabor filters and MLP while the criteria of training stop are fixed to goal=0.005 with learning rate=0.5 and momentum=0.7.

TABLE.3. RECOGNITION RATES USING GABOR FILTERS AND MLP

Number hidden neuron	Recognition Rate (%)
10	94.39
15	94.39
20	94.39
25	94.39
30	94.98
35	94.69
40	93.51
45	94.39
50	94.98

Table 4 gives the recognition rate using the Gabor filters and the constructive training algorithm. We can observe that our approach, with constructive training algorithm, gives the best result, which is substantially higher than the MLP results corresponding to a minimum number of hidden neurons.

Indeed, using the constructive training algorithm, the best recognition rate (97.05%) has been achieved with only 8 hidden neuron and a Eps=0.005. Although, the MLP results has been given a recognition rate equal to 94.98% with 30 hidden neuron.

TABLE.4 RECOGNITION RATES USING THE GABOR FILTERS AND THE CONSTRUCTIVE TRAINING ALGORITHM.

Eps	epochs	N_hid	Recognition rate (%)
0.01	500	9	86.13
0.01	1000	9	91.74
0.01	5000	16	92.92
0.01	10000	23	86.13
0.007	500	45	90.56
0.007	1000	15	92.33
0.007	5000	16	96.16
0.007	10000	47	92.03
0.005	500	178	95.87
0.005	1000	47	94.98
0.005	5000	8	97.05
0.005	10000	12	94.98

Secondly, 5 different scales and 8 different directions are used, $\mu \in \{0,1,2,3,4\}$ and $\nu \in \{0,1,2,3,4,5,6,7,8\}$. Figure 5 shows the magnitude of the Gabor filters at 5 scales and 8 orientations.

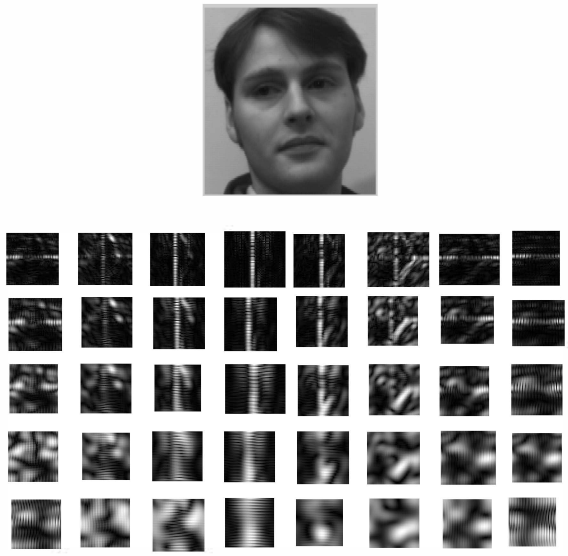

Figure 5 – The original image and the magnitude of the Gabor representation at 5 scales and 8 orientations

Keeping only the magnitude values in the representation, take an image with size 48*48, the convolution result will give 48*48*5*8=92160 features with 5 scales and 8 orientations. It is difficult to calculate with feature vectors of such a high dimension, and a large memory size is also required.

To reduce the dimensionality of this vector we here apply the PCA algorithm to the face Gabor image to extract the feature space.

The results using the GABOR feature and MLP is illustrated in the table 5.

TABLE.5. RECOGNITION RATES USING GABOR FEATURE AND MLP

Number hidden neuron	Recognition Rate (%)
10	92,62 %
15	92,33 %
20	93,33 %
25	93,51 %
30	92,92 %
35	94,39 %
40	93,21 %
45	93,80 %
50	92,92 %

TABLE.6 RECOGNITION RATES USING THE GABOR FEATURE AND THE CONSTRUCTIVE TRAINING ALGORITHM.

Eps	epochs	N_hid	Recognition rate (%)
0.01	500	29	85.84
0.01	1000	18	82.89
0.01	5000	11	80.82
0.01	10000	28	82.89
0.007	500	32	88.79
0.007	1000	29	86.43
0.007	5000	22	96.75
0.007	10000	22	91.74
0.005	500	43	92.92
0.005	1000	32	91.15
0.005	5000	15	96.16
0.005	10000	29	92.33

The recognition rate using the Gabor feature and the constructive training algorithm is shown in table 5.

While comparing results of the MLP and the constructive training algorithm, we can notice that the best recognition rate is gotten with the second architecture with a minimum number of hidden neuron. Indeed, using constructive training algorithm, the best recognition rate (R=96.16%) has been obtained with only 15 hidden neurons and an Eps = 0.005 for a max epochs=5000. But, for the MLP classifier the best recognition rate (94.39 %) has been got with 35 hidden neurons with a goal=0.005.

VI. CONCLUSION

In this paper, a constructive learning algorithm for MLP neural networks has been developed. Starting with a single training pattern and with a one hidden neuron, the proposed algorithm trained incrementally patterns one by one until all patterns of TD are selected and trained to reduce the MSE (converge to a solution). Otherwise, when the training process stuck in a local minimum, the number of hidden neuron added.

The constructive training algorithm for MLP improves the rate of recognition by optimizing the architecture of MLP.

This algorithm is applied in the face recognition system based on the Zernike moment and the Gabor feature. The UMIST database was used for the experiment result. For each feature vector, the best rate recognition obtained using the constructive training algorithm. These results obtained show the effectiveness of the proposed approach.

REFERENCES

[1] W.Zhao and Sarnoff Corporation. "Face Recognition: A Literature Survey". ACM Computing Surveys, Vol. 35, No. 4, December 2003.

[2] M. Turk, and A. Pentland, "Eigenfaces for recognition", Journal of Cognitive Neuroscience, vol. 13, no. 1, pp. 71–86, 1991.

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 19, no. 7, pp. 711–720, 1997.

[4] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: application to face recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 28, no. 12, pp. 2037–2041, 2006.

[5] C. H. The and R. T. Chin, "On Image Analysis by the Methods of Moments, " IEEE Trans. On Pattern Analysis And Machine Intelligence, 10(4), 496-513, 1988.

[6] S.O. Belkasim, M.Shridhar and M. Ahmadi, Pattern Recognition with Moment Invariants: A Comparative Study and New Results, IEEE Trans. On Pattern Recognition, 24(12), 1117-1138, 1991.

[7] A. Khotanzad, and Y.H. Hong, Invariant Image Recognition by Zernike Moments. IEEE Trans. on PAMI, 12 (5). 289-497, 2001.

[8] Javad Haddadnia et-al, "Neural Network Based face Recognition with moment invariants, IEEE transaction, 2001.

[9] C. Liu and H. Wechsler. Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition. IEEE Trans.Image Processing, 11(4) (2002) 467-476.

[10] P.Yang, S. Shan, W.Gao, S.Z.Li, Face recognition using AdaBoosted Gabor features, IEEE international conference on Automatic Face and Gesture Recognition, AFGR'04, 356–361.

[11] V.Kyrki, J.K.Kamarainen, H.Kalviainen. Simple Gabor Feature Space for Invariant Object Recognition. Pattern Recognition Letters, 25(3) (2004) 311-318.

[12] C. Liu, and H. Wechsler, "Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Models for Face Recognition," IEEE Trans, Image Processing, Vol. 11, No. 4, Apr. 2002, pp. 467-476.

[13] Liu, D., Tsu-Shuan, C. and Zhang, Y. (2002) 'A constructive algorithm for feedforward neural networks with incremental training', IEEE Transactions on Circuits and Syst-Fundamental Theory and Applications, Vol. 49, No. 12, pp.1876–1879.

[14] S.Haykin, "Neuronal Networks - A comprehension Foundation," New York, Prentice Hall, 842p,1999.

[15] http://www.sheffield.ac.uk/eee/research/iel/research/face