Volume 3, Issue 5, May 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Genetic Algorithm Based Feature Selection and BPNN Based Classification for Face recognition

Mr. Vivek Pali*
MTech Scholar(DE), RCET, Bhilai India.

Mr. Lalit Kumar Bhaiya
Associate professor, HOD(ET), RCET,Bhilai India.

Abstract— Face Recognition is a progressive field of research which is facing many challenges. This paper proposes a method of genetic algorithm (GA) based neural network for feature selection that retains sufficient information for classification purposes. This method presents an integration of genetic algorithm with an artificial neural network classifier. The proposed system consists of four stages: Eigenface approach or PCA is used for dimensionality reduction, LDA is used for feature extraction, Genetic Algorithm based feature selection and finally Back propagation Neural Network (BPNN) is used for the classification of face images to a particular class. This method uses LDA and PCA for dimensionality reduction and feature extraction which overcomes the small sample size (SSS) problem of LDA. Then among of these extracted features a subset of important features is selected which is applied as an input to the BPNN. The BPNN classifies the face images to a particular class. We have created a train database of 80 images and a test database of 40 images. The proposed system achieved a recognition rate of 97.5% with a small execution time of 10 ms to 20 ms. The results showed that proposed method achieved much better recognition rate than some of the existing face recognition approaches. The experimental results showed significant improvement in execution time required for the face recognition.

Keywords— Genetic Algorithm (GA), Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Back propagation Neural Network (BPNN), Small Sample Size(SSS), Recognition rate.

I. INTRODUCTION

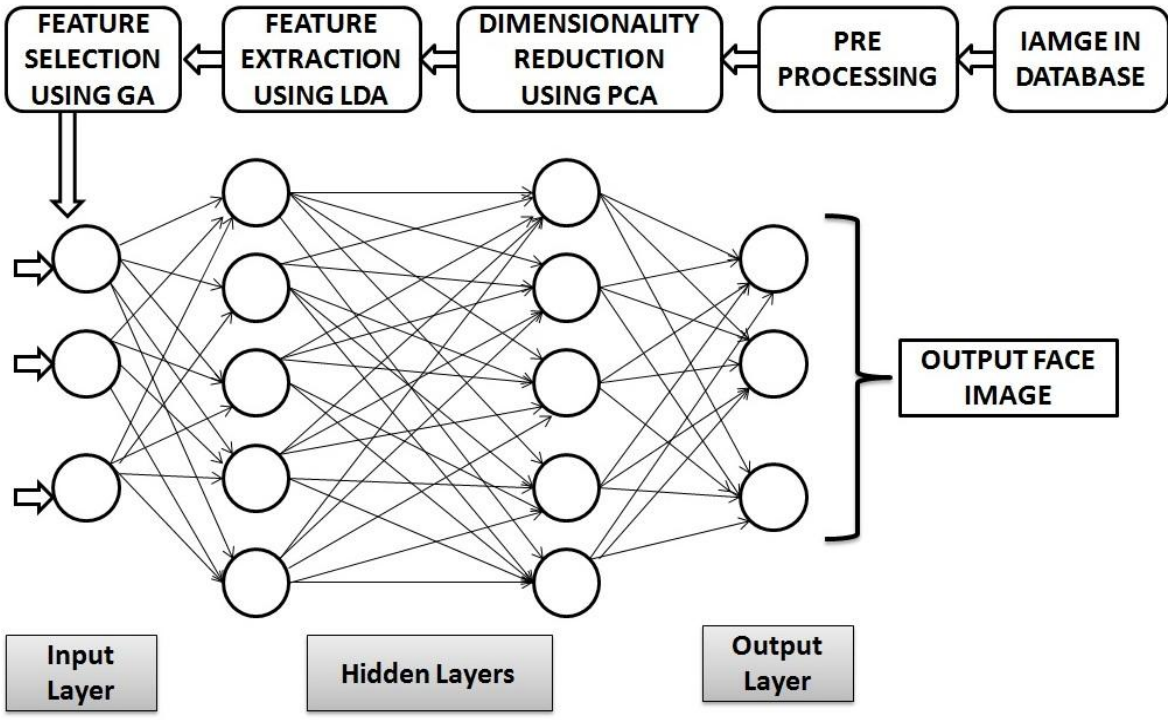
Face recognition has become one of most active research areas of pattern recognition in the last decades. The most of the existing face recognition algorithms can be categorized into three major categories: holistic feature based method; local feature based method and hybrid matching method. In a holistic feature based matching method, the whole face region is used as raw input to the face recognition system. Most popular holistic feature based methods are Principal Component Analysis (PCA) projection method [1], fisherface method [2] and Nearest Feature Line (NFL) method [3]. Recently, a kernel Associative Memory (kAM) models based method [4] and an Independent Gabor Features (IGF) method [5] were developed for the face recognition. In local feature based matching method, local features such as eyes, nose, and mouth are first extracted and then their local statistics (geometric and/or appearance) and locations are fed into a structural classifier. Elastic Bunch Graph Matching (EBGM) method [6] and Geometrical features method [7] falls under this category. For hybrid matching method, both holistic and local features are used for the recognition. A novel feature combination scheme is presented in [8] for face recognition by fusion of global and local features. A fully automatic system for face recognition in databases with only a small number of samples is presented in [9], in which global and local texture features are extracted and used for the classification. Linear discriminant analysis (LDA) and Principal Component Analysis (PCA) both are powerful tools used for dimensionality reduction and feature extraction in face recognition tasks. The major difference between these two methods is that LDA algorithm selects features that are most effective for class separability whereas PCA selects features important for class representation. It has been demonstrated in [10] that when the number of samples per class is small PCA might outperform LDA and the LDA still outperform the PCA, in the case of training set with a large number of samples. Compared to the PCA method, LDA requires much higher computation [11] and PCA is less sensitive to different training data sets. However, simulations reported in [11] demonstrated an improved performance using the LDA method compared to the PCA approach. When dimensionality of face images is high, LDA is not applicable and therefore we derive from its advantage to find effective features for class separability. The classification performance of traditional LDA is often degraded, due to two factors: 1) their classification accuracies suffer from the small sample size problem (SSSP), which widely exists in face recognition; 2) their Fisher discriminant criteria are not directly related to the classification ability. In order to solve the SSSP, which widely exists in FR tasks, a very popular technique usually called PCA plus LDA has been proposed and verified to be effective by experience [11]. In this method, PCA is first utilized for dimensionality reduction before the application of LDA .This paper uses a holistic approach for the face recognition but it combines the two popular approaches PCA and LDA. Genetic Algorithms (GA) has been used to select optimal feature set for pattern classification problems. In this work, GA is used to search the possible optimal subset of extracted features which is then fed to the BPNN classifier which classifies the input test image to a particular class.

II. PROPOSED METHOD

The proposed method is consists of five steps, which excludes face image acquisition for capturing the images to form the database. Firstly, a sequence of pre-processing steps has been done on the input face image. Image resizing, RGB to gray conversion, and Median filtering are included in this pre-processing step. A high dimensional data is associated with this preprocessed image. This high dimensional data is having some useful information which is useful for the classification of face images and in the same time it may have some irrelevant information too. Therefore a dimensionality reduction technique is required to be introduced here. In the proposed work, for dimensionality reduction, Principal Component Analysis (PCA) is used, which reduces this higher dimensional space into lower dimensionality space.

The next step is feature extraction using Linear Discriminant Analysis (LDA). Feature extraction in face recognition problems involves the derivation of salient features from the raw input data in order to reduce dimensionality of facial vectors for classification and simultaneously provide enhanced discriminatory power. LDA has been considered as one of the most effective approaches in the feature extraction of face recognition. Integration of PCA and LDA for feature extraction and automatic feature subset selection distinguishes our proposed facial image classification method from other reported approaches. Facial images are represented in a low dimensional space computed by using PCA+LDA. Now, a GA is used to select facial features automatically and improve the performance of the gender classifier that is implemented using a NN.

Finally, the faces are classified by a neural network technique that is Back Propagation Neural Network (BPNN). Figure 1 shows the process flow of the proposed face recognition system.



.

Figure 1: Process flow of proposed face recognition system

A. *IMAGE ACQUISITION & DATA COLLECTION*

Image acquisition is used to collect the face images; a camera has been used for the face image acquisition. After scanning, the image can be saved into various formats such as Bitmap, JPEG, GIF and TIFF. This system can process face images of any format. By acquiring the images a training database of 80 images and testing database of 40 images with different illumination, poses and expressions is created. The training database contains two images of each person and there are total 80 images of 40 different persons.

B. *IMAGE PRE-PROCESSING*

The next step is image pre-processing which is basically a sequence of a number of steps, which are used to improve the data quality of the image applied to the system and to make the feature extraction phase more reliable. In pattern recognition, the fundamental objective of the image pre-processing phase is to improve the image data quality by suppressing undesired distortions and improving the required image features for further processing. Image preprocessing steps include image resizing, RGB to gray conversion, histogram equalization and noise filtering. The face images stored in the databases must be of same size hence image resizing has to be done on the applied images to resize the images in a standard size. Image resizing here is used to convert the images into a standard size of 180X200. Now, the process of RGB to Gray conversion is done in order to convert this resized image into a single intensity level image (gray image).

Histogram equalization has been applied to adjust the contrast. Finally, Median filtering is done for the noise removal from the images.[12]

C. *DIMENSIONALITY REDUCTION & FEATURE EXTRACTION*

The task of the dimensionality reduction and feature extraction methods is to obtain the most relevant information from the original data and represent that information in a lower dimensionality space. The goal is to transform the origin space X in a new space Y to obtain new features that work better. This way, we can compress the information. Principal Component Analysis (PCA) has been successfully applied to human face recognition. This method is also known as Karhunen-Loeve Transformation or Eigenspace Projection. PCA consists on a transformation from a space of high dimension to another with more reduced dimension. If the data are highly correlated, there is redundant information. PCA decreases the amount of redundant information by decorrelating the input vectors. The input vectors, with high dimension and correlated, can be represented in a lower dimension space and decorrelated. PCA is a powerful tool to compress data. PCA selects features which are important for class representation. The proposed work uses PCA for dimensionality reduction.

Steps to compute the PCA transformation of a data matrix X:

- Center the data $\bar{X}$
- Compute the covariance matrix C
- Obtain the eigenvectors U and eigenvalues ∧ of the covariance matrix
- Project the original data in the eigenspace.

One of the most important components of an automatic face recognition system is the extraction of facial features, in which attempts are made to find the most appropriate representation of face images for identification purposes. The main challenge in feature extraction is to represent the input data in a low-dimensional feature space, in which points corresponding to different poses of the same subject are close to each other and far from points corresponding to instances of other subjects' faces. However, there is a lot of within-class variation that is due to differing facial expressions, head orientations, lighting conditions, etc., which makes the task more complex. Each of the facial components (features) has a different discrimination power for identifying a person or the person's gender, race, and age. LDA extracts these features which are used for classifying the face images. This analysis allows objective evaluation of the significance of visual information in different parts (features) of the face for identifying the human subject. The LDA of faces also provides us with a small set of features that carry the most relevant information for classification purposes. The features are obtained through eigenvector analysis of scatter matrices with the objective of maximizing between-class variations and minimizing within-class variations. (13)

Steps for LDA:-

- There is a conjunct of N vectors of dimension M in the data matrix MxN.
- Assume there C classes and K vectors per class.
- Find the transformation matrix W that better describes the subspace that discriminates between classes, after projecting the data in the new space.
- Calculate between-class scatters S_b & within class scatter matrix S_w
- $S_b = \sum_{j=1}^{C} (\mu j - \mu)\,(\mu j - \mu)^T$
  $S_w = \sum_{j=1}^{C} \sum_{i=1}^{K} (x_i^j - \mu_j)\,(x_i^j - \mu_j)^T$
  The objective is to make maximum the distance between classes S_b and minimizing S_w.
- Calculate the eigenvectors of the projection matrix
  $W=\text{eig}(Sw^{-1}Sb)$
- Compare projection matrices of test images and training images and the result is the training image closest to test image.

D. *FEATURE SELECTION*

There are 4 basic steps in any feature selection method. They are the generation - to generate candidate feature subset, evaluation - to evaluate the generated candidate feature subset and output a relevancy value, where the stopping criteria will determine whether it is the defined optimal feature subset. If yes, the process end else the generation process will start again to generate the next candidate feature subset. Another condition where the process will terminate is when the required number of feature is obtained or the user-specified iteration has reached.

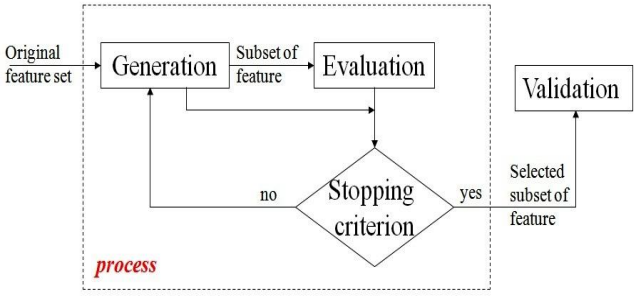


Figure 2: Steps for feature selection process

The last step is validation. This is not a step of the selection process, but a step in the method used to verify that the selected feature subset is a valid subset as required. The proposed work uses GA based optimization technique; this algorithm can efficiently search for the optimal solution, the maximal classification accuracy or minimal classification error rate. For searching a feature space of size *m* chromosomes of length *m* are used. Each bit of every chromosome corresponds to a unique dimension of the feature space and has the value of 1 or 0 and indicates that the corresponding feature is selected or not (Fig. 3).

A1	A2	.....	Am

1 bit for each feature.
Ai= 0 ; the feature is not selected.
Ai= 1 : *the feature is selected.*

Figure 3: Chromosome representation

GA uses genetic operators in order to search the feature space. The genetic operators are selection, crossover and mutation which are driven as a function of a probability distribution by the fitness function. In our experiment, we use
(i) Roulette wheel strategy for selection where each individual is represented by a space that proportionally corresponds to its fitness. By repeatedly spinning the roulette wheel, individuals are chosen to fill the intermediate population.
(ii) Two points crossover where sections between the crossover points are exchange as shown in Fig. 4.
(iii) Fixed probability mutation: each position of a chromosome is given a fixed probability of undergoing mutation (flipping the corresponding bit). In our experiments we used 0.2% as the mutation rate. The first population is generated randomly.

E. *CLASSIFICATION THROUGH BPNN*

Finally, this subset of feature is applied to the BPNN classifier for the classification of face images applied to the neural network. It is a supervised learning method, and is a generalization of the delta rule. It requires a dataset of the desired output for many inputs, making up the training set. It is most useful for feed-forward networks or the networks that have no feedback, or simply, that have no connections that loop. The term is an abbreviation for "backward propagation of errors". Backpropagation requires that the activation function used by the artificial neurons (or "nodes") be differentiable. ANN with Back propagation algorithm is found to be the efficient method for recognizing the faces.

III. RESULT & DISCUSSION

The proposed system is tested using a database containing face images of 40 different subjects. We have two different images of each subject that is taken under different conditions like with various expressions, illumination changes, pose changes etc. Size of each image is 180X200 pixels. The images were taken at different times, varying the illumination, facial expressions such as open or closed eyes, smiling or not smiling etc. and facial details with accessories like sunglasses, no sunglasses, specs, no specs, cap or without cap. The principal components, which are having values close to zero, are omitted to reduce the number of dimensions for computability, while principal component having high eigenvalues are retained for the further processing. The PCA values obtained are then applied to a LDA based feature extraction stage. During this stage, the feature vector is formed using LDA based feature extraction. Now this reduced feature vector is given to the neural network for classification. The rest of the process is managed by the neural network. These feature vectors are applied on a back propagation neural network. Genetic algorithm based feature selection approach significantly improves the accuracy compared to presently available solutions. The fitness of the solutions is improved by applying the mutation and crossover on the population. To get the optimum result this process is repeated several times. The outputs for hybrid model of neural network and genetic algorithm based face recognition system is shown in figure 4 and figure 5.

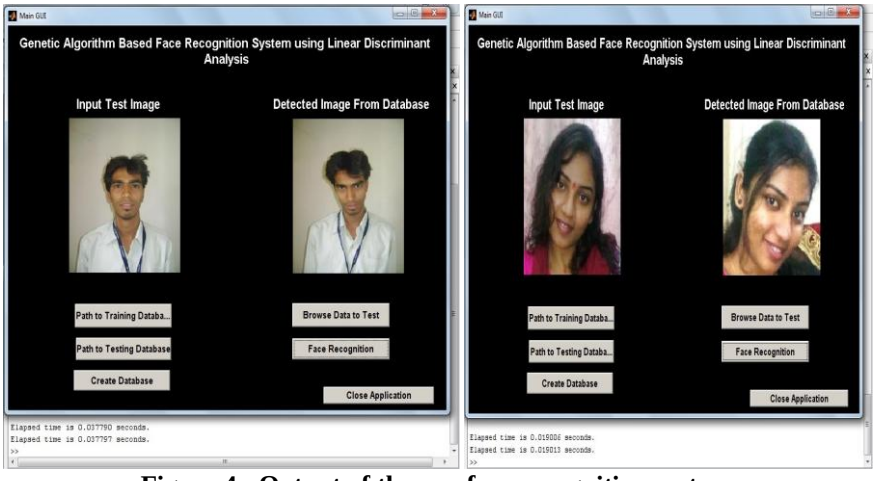


Figure 4 : Output of the our face recognition system

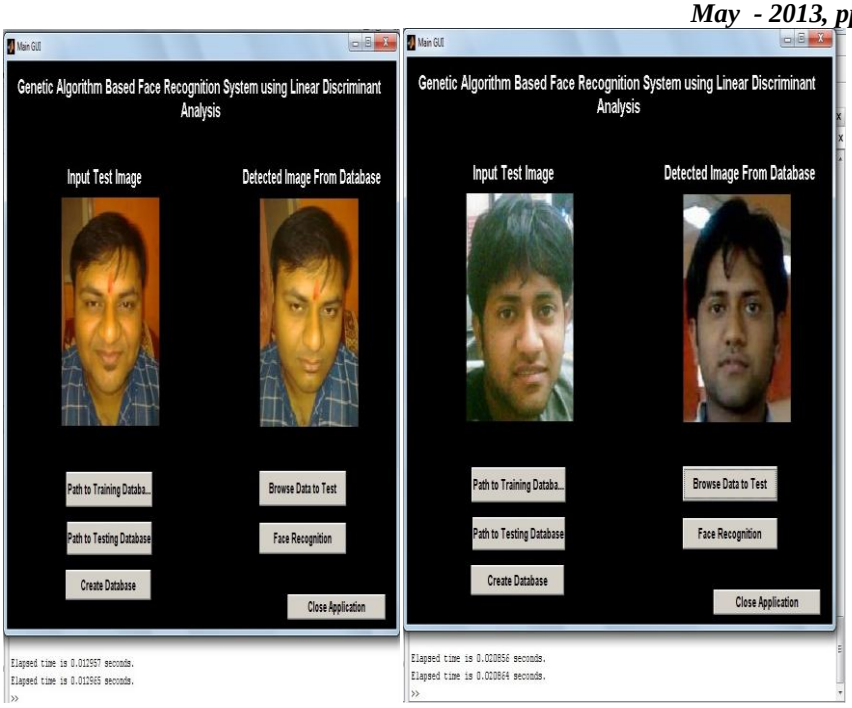


Figure 5 : Output of the our proposed face recognition system

REFERENCES

[1] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, pp.71-86,1991.

[2] 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*, vol. 19, pp.711-720, 1997.

[3] S. Z. Li and J. Lu, "Face recognition using the neareast feature line method," *IEEE Transactions on Neural Networks*, vol. 10, pp. 439-443, 1999.

[4] B. Zhang, H. Zhang and S. Ge, "Face recognition by applying wavelet subband representation and kernel associative memory," *IEEE Transactions on Neural Networks*, vol. 15, pp. 166-177, 2004.

[5] C. Liu and H. Wechsler, "Independent component analysis of Gabor features for face recognition," *IEEE Transactions on Neural Networks*, vol. 14, pp. 919-928, 2003.

[6] L. Wiskott, J. M. Fellous and C. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 775-779, 1997.

[7] R. Brunelli and T. Poggio, "Face recognition through geometrical features," *Proc ECCV92*, pp. 792-800,1992.

[8] Y. Fang, T. Tan and Y. Wang, "Fusion of global and local features for face verification," *IEEE International Conference on Pattern Recognition*, vol. 2, pp. 382–385, 2002.

[9] S. Yan, X. He, Y. Hu, H. Zhang, M. Li and Q. Cheng, "Bayesian shape localization for face recognition using global and local textures," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 102-113, 2004.

[10] A. M. Martinez, A. C. Kak, "PCA versus LDA", IEEE Trans. Pattern Anal. Machine Intell, Vol. 23, pp. 228-233. 2004.

[11] J.J. Weng, "using discriminant eigenfeatures for image retrieval", IEEE Trans. Pattern Anal. Machine Intell. , Vol. 18, No. 8, pp. 831-836, 1996.

[12] Image Processing toolbox TM7 User's Guide.

[13] K. Etemad and R. Chellappa , "Discriminant analysis for recognition of human face images", J. Opt. Soc. Am. A/ Vol. 14, No. 8/August 1997.

