Face Recognition Based on Windowing Technique using DCT, Average Covariance and Artificial Neural Network

Divya A[1], K B Raja[2] and Venugopal K R[3]
[1, 2] Dept. of Electronics and Communication Engineering, University Visvesvaraya College of Engineering, Bangalore University, Bangalore, India
[3] Vice-Chancellor, Bangalore University, Bangalore, India
divyavsagar24@gmail.com[1] and raja_kb@yahoo.com[2]

***Abstract* - The field of Face Recognition (FR) is still a thought-provoking problem, while in recent advances of Artificial Neural Networks (ANN) has shown improved performance in FR rate. In this paper, we propose face recognition based on windowing technique using Discrete Cosine Transform (DCT), average covariance and ANN. The novel concept of windowing technique is used to divide each image to 4X4, 8X8 and 16X16 size of windows. The DCT is applied on each window to obtain DCT co-efficients. The covariance matrix is computed on each DCT coefficient matrix and average value of each block is also computed to obtain final feature value. The computation of an average covariance reduces the original size of face image by around 97% i.e., the number of co-efficients in the final feature set is only around 3% of the original size of an image. The proposed method is very efficient in identifying with very less number of features. Network is created and trained the input dataset and target dataset to reach the desired output. The trained net is then tested to compute performance parameters of the network. The experiments are conducted on some popularly used face databases to illuminate the performance and the efficiency of the proposed algorithm. The experimental results are tabulated and are compared with the existing methods. It is observed that, the proposed model achieves better recognition accuracy for 16X16 windowing and also with existing algorithms.**

***Keywords*—Biometrics; Discrete Cosine Transform (DCT); convergence; Artificial Neural Networks (ANN).**

I. Introduction

Face Recognition is an evolving biometric modality that has been fascinated a lot of research activities in the field of image processing and computer vision for its increased accuracy, ease-of–use and reduced cost. FR is widely used in the field of social networking, video surveillance, biometric passwords. Many effective feature extraction techniques like Local Binary Pattern (LBP) [1], DCT [2], Discrete Wavelet Transform (DWT) [3] etc., have been proposed to analyze facial expressions. Several feature extraction techniques have presented that the features learned from training the network offers a significantly enhanced the performance of recognition related to the conventional methods of feature extractions for facial expressions.

An Artificial Neural Network (ANN) [4] based approaches has become more popular due to its robustness of comprehensive feature extraction from the face data. It has shown robust recognition techniques to the intra-class spatial variations and has wide range of applications like character recognition, object detection, image classification and face recognition. The ANN has significantly improved the performance in the facial expression recognition compared to the traditional methods. This motivates to explore ANN to bring about enhanced performance for the many face recognition problems. The ANN models has different types of networks like Learning Vector Quantization (LVQNet), Feed Forward Back Propagation Network (FFBPNN), FitNet, Cascade Forward Back Propagation Network (CFBPNN) [5] and the most popular ANN model is the multi-layer feed forward networks FFBPNN which use backpropagation algorithm for training.

The key contribution of this paper is to exhibit that, the Average Covariance of DCT for feature extraction is an effective way to construct an index for neural network classifier. The covariance of DCT co-efficients that are obtained for the 4X4, 8X8 and 16X16 windowing and averaged to obtain the final features of face images. The final features of training dataset are fed as input to the NN and trained using target. The output of NN is tested for the performance of the proposed method.

The paper is structured as follows: Section II briefly reviews the existing techniques of face recognition and deep learning. In section III, background work is explained in detail. The proposed model for face recognition scheme is illustrated in Section IV. An extensive theoretical and experimental outcomes and the comparison of proposed methods with the conventional methods are presented in Section V, leading to conclusion in Section VI.

II. Related Work

In this section, several popular feature extraction techniques of biometrics and deep learning are discussed. Mahbubul Alam et al., [6] proposed a biological related Sparse-deep Simultaneous Recurrent Network(S-DSRN) for the facial expression recognition. The proposed method proposes the desirable properties such as sparsity and also overcomes from overfitting problem. It provides an efficient control over the depth model by having constant number of training parameters

978-1-5386-7516-8/18/$31.00 ©2018 IEEE

by sharing the weights between the hidden layers. Experiment was conducted for two widely used dataset and resulted in better recognition accuracy by reducing the computational complexity with fewer number of parameters. The results also show that the integration of DML-Eigen for the proposed S-DSRN model offers a considerable improved performance over common regression based classification like softmax. Tong Zang et al., [7] proposed a Deep Neural Network (DNN) based feature learning method which is used for Facial Expression Recognition (FER). The Scale Invariant Feature Transform (SIFT) is used to extract the essential features of face images and the obtained feature vectors are fed into a well-made DNN model for optimum learning of maximum features for facial expression classification. The proposed method consists of number of layers which are trained and are able to learn a set of optimum features, which helps for classifying the facial expressions. The performance evaluation is performed on BU-3DFE and Multi-PIE face databases. The experimental results indicate that the proposed model performs better compared to the traditional methods. Zijing Zhao and Ajay Kumar [8], proposed a Semantics-assisted Convolutional Neural Networks (SCNN) which includes the obvious semantic information to get better comprehensive features. The proposed method is robust, more accurate framework and is capable of extracting more features from the face images and thereby achieving a very good performance. The experiment was performed on four databases and results achieved outstanding performance with smaller computational time for matching process. Kaihao Zhang et al., [9] proposed a Part-based Hierarchical bidirectional Recurrent Neural Network (PHRNN) and Multi Signal CNN (MSCNN) to analyze the facial expression information of temporal sequence and spatial feature from still frames respectively. The proposed methods extract whole geometrical and dynamic-still information, resulting in outstanding performance compared to the traditional ones. Fanlong Zhang et al., [10] presented a structured 2D method known as Nuclear norm-based 2DPCA (N-2DPCA) and a Bilateral projection based N-B2-DPCA is also proposed. The benefit of N-B2-DPCA over N-DPCA is that an image can be represented with lesser co-efficient values and can be applied to face recognition and as well as for face reconstruction. Performance evaluation is performed on various face databases such as Extended Yale B, CMU PIE, FRGC and AR. The experiments are conducted and results shows the effectiveness of the proposed method over the conventional one.

III. Background

This section delivers a brief summary on face image databases, DCT and ANN methods.

A. Face Image Databases

The publicly used face databases such as ORL, YALE, JAFFE, Indian Male, Indian Female, NIR and Extended Yale B are used to test the performance of the proposed model.

1) ORL Face Database [11]: The Olivetti Research Laboratory (ORL) database has total four hundred face images of forty subjects with ten images per subject, with the image of size 112X92. All the images are taken in contradiction of the dark background in frontal position. The captured images are with some specifications with occlussions and different lighting conditions, facial expressions, different face angles and facial details.

2) YALE Face Database[12]: The database includes 150 face images in total with 15 subjects and 10 pictures per subject. Each image is of size 243X320 with jpg format. The images are captured against the white background with different lighting angle conditions. The face images are captured with different facial expressions.

3) JAFFE Face Database[13]: The Japanese Female Facial Expressions (JAFFE) database has 10 subjects and per subject 23 images, therefore the total number of images are 230. The size of the image is 256X256 with .tiff format. The face images are taken for various expressions like anger, smiling, etc.

4) Indian Male Face Database[14]: The face database consists of 20 subjects with 11 images per subject i.e., 220 images in total. The images are of size 480X640 with jpg format. The RGB images are captured under dissimilar facial orientations and expressions.

5) Indian Female Face Database[14]: The face database has 22 subjects with 11 images per subjects and the total number of images are 242. Each size of the image is 480X640 with the jpg format and are RGB images. The face images are taken with different facial angles and expressions.

6) NIR Face Database[15]: The Near Infrared database has 120 subjects with 15 images per subject. Total number of images are 1800 and each image is of size 576X768. The images are captured under the dark background with changed lighting conditions and facial expressions.

7) Extended Yale B Face Database[16]: The Extended Yale B (EYB) face database contains 2141 face images of 38 subjects. The face images are captured under altered illumination circumstances. All the images are in cropped version and are resized to the size of 192X168.

B. Discrete Cosine Transform (DCT)

The Transform is associated to the Discrete Fourier Transform (DFT). DCT is distinguishable, linear transformation and are mainly used for image compression applications.

The definition for 2D-DCT for an input image *X* and output image *Y* is given by the following Eq. (1), (2) and (3).

$$Y_{pq} = \alpha_p \, \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{mn} \cos\frac{\pi(2m+1)p}{2M} \cos\frac{\pi(2n+1)q}{2N}\,, \tag{1}$$

where $0 \le p \le M-1$; $0 \le q \le N-1$

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & 1 \le p \le M-1 \end{cases} \tag{2}$$

$$\alpha_q = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & 1 \leq q \leq M-1 \end{cases} \qquad (3)$$

where M = row size of X
N = column size of X.

C. Covariance

It is a linear statistical measure of dependence. For a matrix A whose columns and rows represents random variable and observations respectively. The covariance matrix C for a matrix A is given by below Eq. (4).

$$C(i,j) = \mathrm{cov}(A(:,i), A(:,j)) \qquad (4)$$

where i = rows of the matrix A
j = columns of the matrix A.

D. Artificial Neural Network (ANN)

It basically consists of P number of inputs which is multiplied by the weights w and then passed through a mathematical function resulting in an output a. The Neural Network (NN) generally be made up of three layers viz., input layer, hidden layer and the output layer.

ANN process the data one at a time, and learn the network by relating their classification of the data with the known actual classification of the data. The errors generated from the early classification is fed back into the network to adapt the network algorithm for second time classification, and this process is done for many iterations to get the desired output. There are various types of learning algorithms which are available for the ANN and Delta rule is one among them which is most commonly used algorithm by the classes of ANN called Backpropagation Neural Network (BPNN).

The training data consists of input data and target data, which are fed to the input layer during the training phase. The hidden layer acquires the input from all the nodes of the input layer, the weights are initialized and then multiplied with each of the input values. The NN is trained on a representative set of input and target datasets to obtain the desired output. The output layers lead to a set of real output values, which is related with the target sets. The errors generated among the target and actual output values are computed using the sum of square errors and propagated back to the network, called backward pass. The backpropagation neural network algorithm optimizes the system by choosing the best weight co-efficients to minimize the errors. The test data set and validation set is not trained during the training phase. The validation dataset is used to validate the trained network whereas the test dataset is used to check how well the network is learned during the training phase and to calculate the performance of the network.

IV. PROPOSED MODEL

In this section, the proposed face recognition model based on windowing technique, DCT and average covariance using Feed Forward Backpropagation Neural Network (FFBPNN) classification is implemented and discussed in detail and the block diagram is as shown in the Fig. 1.

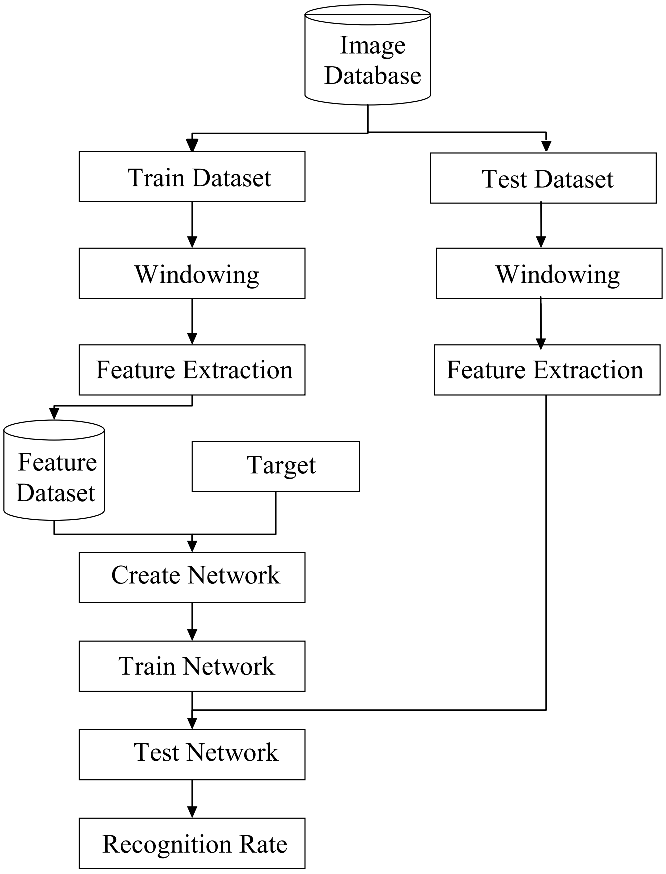


Fig. 1. Block diagram of face recognition system using proposed model.

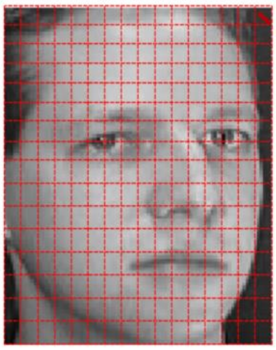
A. Face Image Database

The face databases such as ORL, Yale, JAFFE, Indian Male, Indian Female, NIR and EYB having various sizes, illuminations, poses and background are considered to test the proposed method.

B. Windowing Technique

The face images are considered from the various standard face databases and are pre-processed by converting some of the RGB images into grayscale images in order to reduce the computations.

For the illustration of the proposed model, the ORL database is considered and size of each image is 112X92 with pgm format. The face images in the database are divided into windows that are of sizes 4X4, 8X8 and 16X16. Each cells are of sizes 28X23, 14X11 and 7X5 for the 4X4, 8X8 and 16X16 windows respectively for original size of 112X92. The 16X16 window with cell size of 7X5 matrix is considered for demonstration purpose with the pixel values, shown in Fig. 2.



$$\begin{bmatrix} 60 & 60 & 62 & 53 & 48 \\ 58 & 73 & 54 & 63 & 62 \\ 68 & 77 & 80 & 67 & 57 \\ 79 & 76 & 56 & 62 & 64 \\ 49 & 69 & 68 & 63 & 79 \\ 43 & 65 & 89 & 79 & 68 \\ 52 & 62 & 67 & 71 & 82 \end{bmatrix}_{7X5}$$

Fig. 2. Pixel values of 7X5 for 16X16 window size cell

C. *Feature Extraction*

1) DCT2 on each cell of windowing Image: The DCT is applied on each of the cells of size 4X4, 8X8 and 16X16 to obtain the corresponding DCT co-efficients. The DCT co-efficients for 16X16 cell of Fig.2 is as shown in Fig. 3.

$$\begin{bmatrix} 386.2355 & -8.2225 & -15.1651 & -12.6222 & -3.8104 \\ -16.0620 & 31.3364 & 5.9208 & 3.3276 & -5.5680 \\ -13.7872 & -13.6020 & -5.7534 & 3.4427 & 3.8004 \\ -10.8993 & -5.5820 & -3.1918 & -2.6484 & 4.1582 \\ -4.3475 & 13.5171 & 18.9555 & 2.6682 & 0.6123 \\ 6.5589 & 6.9342 & -20.9887 & 10.2078 & 9.3079 \\ 0.5950 & -5.3113 & -6.1824 & -4.3775 & 7.6711 \end{bmatrix}_{7X5}$$

Fig.3. DCT co-efficients of a cell for window size of 16X16

2) Covariance of the DCT co-effiecients: The covariance of Fig. 3 is computed and resulting in 5X5 matrix as shown in Fig. 4.

$$\begin{bmatrix} 22079.54 & -732.67 & -791.37 & -815.72 & -370.28 \\ -732.67 & 245.46 & 103.62 & 51.02 & -38.33 \\ -791.37 & 103.62 & 174.46 & 10.52 & -30.08 \\ -815.72 & 51.02 & 10.52 & 53.29 & 15.78 \\ -370.28 & -38.33 & -30.08 & 15.78 & 30.95 \end{bmatrix}_{5X5}$$

Fig.4. covariance matrix of DCT co-efficients

3) Average Covariance: The average of covariance matrix of Fig. 4 is computed and resulting in one value of 695.5475. Similarly average covariance for all 16X16 blocks of an image are computed to obtain 16X16 average covariance matrix as shown in Fig. 5.

$$\begin{bmatrix} 695.5475 & 999.3077 & \cdots\cdots\cdots & 636.5961 & 173.1475 \\ 1157.985 & 1873.865 & & 6322.176 & 6208.3 \\ & \vdots & & \vdots & \\ & \vdots & & \vdots & \\ 4062.483 & 5408.481 & & 4387.94 & 3689.533 \\ 2557.518 & 4810.161 & \cdots\cdots\cdots & 3124.22 & 3398.505 \end{bmatrix}_{16X16}$$

Fig. 5. Average covariance values of each cell

The original face image size of 112X92 i.e., 10,304 pixels are converted into 16X16 average covariance matrix with 256 co-efficients. The total number of pixels in the original image are compressed by around 97% in the average covariance matrix. Hence the number of co-efficients in the final feature set has only around 3% of original size of an image, which results in fast computation in real time applications. The final feature set matrix of size 16X16 is converted into column vector of size 256X1 as shown in Fig. 6 and is applied as input to NN for effective identification of face image.

$$\begin{bmatrix} 695.5475 \\ 1157.9817 \\ 1904.9783 \\ 2710.4499 \\ 3734.4885 \\ \vdots \\ \vdots \\ \vdots \\ 4580.5129 \\ 4663.7542 \\ 3921.9287 \\ 3689.5328 \\ 3398.5047 \end{bmatrix}_{256X1}$$

Fig. 6. Feature vector for one image

D. *Classification using Neural Networks*

The final features values for the dataset are fed as input to the NN. The number of values for each final feature image of ORL dataset are 16, 64 and 256 for 4X4, 8X8 and 16X16 windows respectively. The total number of images in the ORL database is equal to four hundred and all the images are divided into training set, validation set and testing set with percentage combinations of 70:15:15 i.e., 280 images for training, 60 images for validation and 60 images for testing respectively. The total number of images are also divided into new combination of 80:10:10 i.e., 320 images for training, 40 images for validation and 40 images for testing respectively. The size of final feature set for 280 training images are 16X280, 64X280 and 256X280 for window sizes of 4X4, 8X8 and 16X16 respectively. Similarly, the size of final feature set for 320 training images are 16X320, 64X320 and 256X320 respectively. The target data for pattern recognition network is 40X280 matrix i.e., the number of classes (persons) are 40 and total number of training images are 280. All the elements of matrix have zero values except the element belong to corresponding class and assign the value one.

The target data for pattern recognition network is a 40X280 matrix, consists of all zero values except for the elements belonging to corresponding class. The pattern recognition network is created using the number of hidden layers, transfer function and error function, the network is trained using trainscg. The network is trained constantly for number of iterations. Throughout the training phase, the network output is compared with the target data and the proper error is calculated. The obtained error value is redistributed back to the hidden layer, for updating the weights consequently and the network is trained to reach the desired output values. The network is tested against the fresh data to evaluate the network and to check the performance of the network. If the performance of the network is not satisfied on the new data, then either train the network again or increase the number of neurons. If the training set performance is high and the test set performance is significantly low, then this could specify the overfitting problem. By decreasing the number of neurons, overfitting problem can be minimized and can improve the test performance. The performance parameters are obtained from confusion matrix.

V. Experimental Results and Discussion

In this section, the definitions of performance parameters and their evaluations for face recognition are discussed. The performance of the proposed model is assessed on seven publicly available face databases.

A. *Definitions of Performance Parameters*

The performance parameters are based on four outcomes – True Positive (TP) - correct positive prediction, True Negative (TN) - correct negative prediction, False Positive (FP) - incorrect positive prediction and False Negative (FN) - incorrect negative prediction. The performance measures are defined as follows:

1) Accuracy (ACC) or Recognition Rate: The ratio of true values to the combination of true and false values, as given in Eq. 5.

$$\% \text{ ACC} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} * 100 \tag{5}$$

2) Average Cross Entropy Error (ACE): It is the loss function used in the machine learning and is given in the Eq. 6.

$$\text{ACE} = 1/N \ \Sigma^{N}_{n=1} H(P_n, Q_n) \tag{6}$$

B. *Performance Evaluation using the standard Face databases*

The efficiency of the proposed model is evaluated using face databases viz., ORL, Yale, JAFFE, Indian Male, Indian Female, NIR and Extended Yale B. The performance parameters such as overall accuracy, test accuracy and average cross entropy error are computed.

1) Results using ORL Face Database: All four hundred face images of ORL face database are used to test the proposed model. The face images are segmented into 4X4, 8X8 and 16X16 sizes using windowing technique. The number of images used for training, validation and test are in the ratio of 70:15:15 and 80:10:10 from the total of 400 images, i.e., 280:60:60 and 320:40:40. The number of hidden layers used in the neural network are twenty-five. The performance parameters such as overall accuracy, test accuracy and average cross-entropy error tabulated in the Table I using window sizes of 4X4, 8X8 and 16X16 for ORL database.

TABLE I
PERFORMANCE PARAMETERS FOR VARIOUS WINDOW SIZES ON ORL DATABASE

Window Size	No. of Training images	Overall Accuracy (%)	Test Accuracy (%)	ACE
4X4	280	97.00	98.7695	0.0028
4X4	320	98.25	98.5913	0.0015
8X8	280	99.00	99.7085	0.0010
8X8	320	99.25	99.8074	0.00044
16X16	280	98.75	99.8123	0.00085
16X16	320	99.75	99.8707	0.00017

The results for ORL face database indicates that the value of overall accuracy and testing accuracy improve with an increase in number of training images. The accuracy also increases with increase in window size. It is also observed that the error decreases with increases in number of training images and size of the window.

2) Results using Yale Database: All one hundred and fifty face images of Yale face database are used to test the proposed method. The face images are segmented into 4X4, 8X8 and 16X16 sizes using windowing technique. The number of images used for training, validation and test are in the ratio of 70:15:15 and 80:10:10 from the total of 150 images, i.e., 104:23:23 and 120:15:15. The number of hidden layers used in the neural network are twenty-five. The performance parameters such as overall accuracy, test accuracy and average cross-entropy error tabulated in the Table II using window sizes of 4X4, 8X8 and 16X16.

TABLE II
PERFORMANCE PARAMETERS FOR VARIOUS WINDOW SIZES ON YALE DATABASE

Window Size	No. of Training images	Overall Accuracy (%)	Test Accuracy (%)	ACE
4X4	104	68.00	92.0614	0.0745
4X4	120	93.333	98.2523	0.0200
8X8	104	92.667	93.8698	0.0249
8X8	120	95.333	94.8820	0.0156
16X16	104	97.333	96.8677	0.0072
16X16	120	99.333	99.6076	0.0015

It is observed that the performance of the overall accuracy and testing accuracy is improved with an increase in number of training images and window size for Yale face database. It is also observed that the error decreases with increases in number of training images and with increase in size of the window.

3) Results using JAFFE Database: All two hundred face images of JAFFE face database are used to test the proposed model. The face images are segmented into 4X4, 8X8 and 16X16 sizes using windowing technique. The number of images used for training, validation and test are in the ratio of 70:15:15 and 80:10:10 from the total of 200 images, i.e., 140:30:30 and 160:20:20. The number of hidden layers used in the neural network are twenty-five. The performance parameters such as overall accuracy, test accuracy and average cross-entropy error tabulated in the Table III using window sizes of 4X4, 8X8 and 16X16.

TABLE III
PERFORMANCE PARAMETERS FOR VARIOUS WINDOW SIZES ON JAFFE DATABASE

Window Size	No. of Training images	Overall Accuracy (%)	Test Accuracy (%)	ACE
4X4	140	100	99.9093	0.0009
4X4	160	99.5	99.9322	0.0011
8X8	140	100	99.9959	0.000007
8X8	160	100	99.9964	0.000004
16X16	140	100	99.9997	0.000003
16X16	160	100	99.9999	0.0000006

It is observed from the results that, the value of overall accuracy and testing accuracy improve with an increase in number of training images and window size. It is also observed that the error is significantly decreased with increase in number of training images and as well as size of the window.

4) Results using Indian Male Database: All two hundred face images of Indian Male face database are used to test the proposed model. The face images are segmented into 4X4, 8X8 and 16X16 sizes using windowing technique. The number of images used for training, validation and test are in the ratio of 70:15:15 and 80:10:10 from the total of 200 images, i.e., 140:30:30 and 160:20:20. The number of hidden layers used in the neural network are twenty-five. The performance parameters such as overall accuracy, test accuracy and average cross-entropy error tabulated in the Table IV using window sizes of 4X4, 8X8 and 16X16.

TABLE IV
PERFORMANCE PARAMETERS FOR VARIOUS WINDOW SIZES ON INDIAN MALE DATABASE

Window Size	No. of Training images	Overall Accuracy (%)	Test Accuracy (%)	ACE
4X4	140	83.00	95.9158	0.0290
4X4	160	83.00	95.8908	0.0287
8X8	140	93.50	97.4543	0.0165
8X8	160	95.50	97.8468	0.0145
16X16	140	97.00	97.1920	0.0182
16X16	160	98.50	98.8577	0.0056

It is observed that the performance of the overall accuracy and testing accuracy is improved with an increase in number of training images and window size for Indian Male face database. It is also observed that the error decreases with increases in number of training images and with increase in size of the window.

5) Results using Indian Female Database: All two hundred and forty-two face images of Indian Female face database are used to test the proposed model. The face images are segmented into 4X4, 8X8 and 16X16 sizes using windowing technique. The number of images used for training, validation and test are in the ratio of 70:15:15 and 80:10:10 from the total of 242 images, i.e., 170:36:36 and 194:24:24. The number of hidden layers used in the neural network are twenty-five. The performance parameters such as overall accuracy, test accuracy and average cross-entropy error tabulated in the Table V using window sizes of 4X4, 8X8 and 16X16.

TABLE V
PERFORMANCE PARAMETERS FOR VARIOUS WINDOW SIZES ON INDIAN FEMALE DATABASE

Window Size	No. of Training images	Overall Accuracy (%)	Test Accuracy (%)	ACE
4X4	170	94.6280	97.3925	0.0103
4X4	194	96.6942	96.7933	0.0088
8X8	170	97.1074	98.5623	0.0049
8X8	194	97.9338	98.2079	0.0053
16X16	170	96.6942	99.1873	0.0077
16X16	194	98.7600	99.3067	0.0036

The recognition results indicate that the value of overall accuracy and testing accuracy improve with an increase in number of training images. The accuracy also increases with increase in window size. It is also observed that the error decreases with increases in number of training images and size of the window.

6) Results using NIR Database: All one thousand eight hundred face images of NIR face database are used to test the proposed model. The face images are segmented into 4X4, 8X8 and 16X16 sizes using windowing technique. The number of images used for training, validation and test are in the ratio of 70:15:15 and 80:10:10 from the total of 1800 images, i.e., 1260:270:270 and 1440:180:180. The number of hidden layers used in the neural network are twenty-five. The performance parameters such as overall accuracy, test accuracy and average cross-entropy error tabulated in the Table VI using window sizes of 4X4, 8X8 and 16X16.

The results indicates that the value of overall accuracy and testing accuracy significantly improve with an increase in number of training images. The accuracy also increases with increase in window size. It is also observed that the error decreases with increases in number of training images and size of the window.

TABLE VI

PERFORMANCE PARAMETERS FOR VARIOUS WINDOW SIZES ON NIR DATABASE

Window Size	No. of Training images	Overall Accuracy (%)	Test Accuracy (%)	ACE
4X4	1260	87.611	99.2557	0.0056
4X4	1440	86.666	99.2088	0.0060
8X8	1260	89.2222	99.3489	0.0044
8X8	1440	88.1667	99.3923	0.0047
16X16	1260	94.4444	99.4466	0.0026
16X16	1440	96.9444	99.4130	0.0015

7) Results using Extended Yale B: All one thousand nine hundred face images of EYB face database are used to test the proposed model. The face images are segmented into 4X4, 8X8 and 16X16 sizes using windowing technique. The number of images used for training, validation and test are in the ratio of 70:15:15 and 80:10:10 from the total of 1900 images, i.e., 1330:285:285 and 1520:190:190. The number of hidden layers used in the neural network are twenty-five. The performance parameters such as overall accuracy, test accuracy and average cross-entropy error tabulated in the Table VII using window sizes of 4X4, 8X8 and 16X16.

TABLE VII
PERFORMANCE PARAMETERS FOR VARIOUS WINDOW SIZES ON EYB DATABASE

Window Size	No. of Training images	Overall Accuracy (%)	Test Accuracy (%)	ACE
4X4	1330	60.5263	95.5598	0.0384
4X4	1520	63.5789	95.7814	0.0342
8X8	1330	82.8947	97.6159	0.0193
8X8	1520	87.5789	97.9249	0.0138
16X16	1330	94.8947	98.6015	0.0055
16X16	1520	95.4736	98.6113	0.0059

The results indicate that the value of overall accuracy and testing accuracy improve with an increase in number of training images. The accuracy also increases with increase in window size. It is also observed that the error decreases with increases in number of training images and size of the window.

The performance of the proposed method with windowing size of 16X16 is better compared to 8X8 and 4X4 windowing sizes for seven standard face databases

C. Comparison of Proposed Method with the Existing Methods

The experimental results shown in Table VIII demonstrates the comparison of our proposed method with the other popular methods for different face databases.

It shows that the proposed method with 16X16 windowing size attains considerably higher recognition rate compared with other existing algorithms for ORL and Yale databases. It is observed that the accuracy of the proposed method outperforms than the popular methods on JAFFE database and it is also observed that, the proposed model achieves as good as state-of-the-art methods on EYB database.

TABLE VIII
COMPARISON WITH OTHER POPULAR METHODS ON THE PUBLICLY AVAILABALE DATABAES

Face Databases	Algorithms	Accuracy (%)
ORL	Windowing average [17]	96
	Decomposed averaged approximation co-efficients [18]	97
	Proposed Method	99.75
YALE	HTM Spatial Pooler [19]	86.67
	RFDCNN [20]	96.47
	Proposed Method	99.33
JAFFE	Decomposed averaged approximation co-efficients [18]	97
	DCNN(Ten-fold) [21]	98.12
	Proposed Method	100
EYB	PNNBIM [22]	95.10
	LLKc [23]	95.39
	Proposed Method	95.47

The comparison of experimental results implicates that the proposed work attains greater recognition rate compared with the existing algorithms in the area of face recognition is due to the following reasons.

(i)The novel concept of windowing technique with larger windowing size is used to divide the image and applying DCT, covariance, averaging on each window cell resulting in lesser feature values.

(ii) Feed Forward Backpropagation Neural Networks are considered to train and test the input dataset.

(iii) Larger training images are considered.

(iv) The proposed model is applied to both smaller and larger datasets, resulting in significantly higher recognition accuracy for both smaller and larger datasets.

VI. CONCLUSION

This paper presents a novel face recognition based windowing technique using DCT and average covariance with neural networks as a classifier results in less number of final feature co-efficients. The standard face databases are used to evaluate the performance of face recognition system. The face images of all the subjects are divided into cells 4X4, 8X8 and 16X16 resulting in each cell size of 27X23,14X11 and 7X5 respectively for ORL face image of size 112X92. The DCT are applied and covariance is computed for the DCT co-efficients and averaging is performed for each cell. The resulting feature vector per image is 16X1, 64X1 and 256X1 for the cell size of 4X4, 8X8 and 16X16 respectively. The obtained feature vector is given as input for the neural network system with corresponding target values. The network is created and trained to get the desired output. Finally, network is tested and resulted in higher recognition rate with 16X16 window compared to 4X4 and 8X8. It is also observed that, the recognition rate of the proposed method is higher compared to the existing methods. In future, different feature extraction techniques and classifiers with minimum hidden layers can be used to increase the performance of the method.

REFERENCES

[1] Juan.E.Tapia and Claudio.A.Perez,"Gender Classification Based on Fusion of Different Spatial Scale Features Selected by Mutual Information from Histogram of LBP, Intensity, and Shape," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 3, Mar. 2013.

[2] Mehran Kafai, Kave Eshghi and Bir Bhanu,"Discrete Cosine Transform Locality-Sensitive Hashes for Face Retrieval," *IEEE Transactions on Multimedia*, vol. 16, no. 4, pp. 1090-1103, June 2014.

[3] Cheolkon Jung and Jianpeng Yin,"SQI-based Illumination Normalization for Face Recognition Based on Discrete Wavelet Transform," *IEEE Conference on Image Processing*, pp. 1664 - 1668, Sept. 2016.

[4] Manisha M. Kasar, Debnath Bhattacharyya and Tai-hoon Kim,"Face Recognition Using Neural Network: A Review," *International Journal of Security and Its Applications*, vol. 10, no. 3, pp. 81-100, 2016.

[5] Omaima N. Ahmad AL-Allaf, Shahlla A. AbdAlKader and Abdelfatah Aref Tamimi, "Pattern Recognition Neural Network for Improving the Performance of Iris Recognition System," *International Journal of Scientific & Engineering Research*, vol. 4, issue 6, pp. 661-667, June 2013.

[6] Mahbubul, Alam, Lasitha S. Vidyaratne, and Khan M. Iftekharuddin, "Sparse Simultaneous Recurrent Deep Learning for Robust Facial Expression Recognition," *IEEE Transactions on Neural Networks and Learning Systems*, vol. PP, issue 99, pp. 1-12, 2018.

[7] Tong Zhang, Wenming Zheng, Zhen Cui, Yuan Zong, Jingwei Yan and Keyu Yan, "A Deep Neural Network Driven Feature Learning Method for Multi-view Facial Expression Recognition,"*IEEE Transactions on Multimedia*, vol. 18, issue 12, pp. 2528-2536, Dec 2016.

[8] Zijing Zhao and Ajay Kumar, "Accurate Periocular Recognition under Less Constrained Environment Using Semantics-Assisted Convolutional Neural Network," *IEEE Transactions on Information Forensics and Security*,vol.12, issue 5, pp. 1017-1030, May 2017.

[9] Kaihao Zhang, Yongzhen Huang, Yong Du and Liang Wang, "Facial Expression Recognition Based on Deep Evolutional Spatial-Temporal Networks," *IEEE Transactions on Image Processing*, vol. 26, issue 9, pp. 4193-4203, Sept 2017.

[10] Fanlong Zhang, Jian Yang, Jianjun Qian and Yong Xu, "Nuclear Norm-Based 2-DPCA for Extracting Features From Images," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, issue 10, pp. 2247- 2260, Oct. 2015.

[11] AT&T Laboratories Cambridge, "The ORL Database of Faces,"1994.[Online].Available:http://www.cl.cam.ac.uk /research/dtg/attarchive/facedatabase.html. [Accessed: 07- July-2017].

[12] Yale University, "The Yale Face Database," 1997. [Online].Available:http://cvc.cs.yale.edu/cvc/projects/yalefaces/yalefaces.html. [Accessed: 30- July-2017].

[13] IIT Kanpur campus, "Indian Face Database," 2002. [Online]. Available: http://vis-www.cs.umass.edu/~vidit/IndianFaceDatabase/. [Accessed: 15- July-2017].

[14] Michael J. Lyons, "The Japanese Female Facial Expression(JAFFE)Database,"1998.[Online].Available:http://www.kasrl.org/jaffe.html.[Accessed:15- June-2017].

[15] The Biometric Research Centre (UGC/CRC), "The Hong Kong Polytechnic University (PolyU) NIR Face Database,"2009.[Online].Available:http://www4.comp.polyu.edu.hk/~biometrics/polyudb_face.htm. [Accessed: 07- July-2017].

[16] Georghiades, A.S. and Belhumeur, P.N. and Kriegman, D.J, "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose," *IEEE Transaction on Pattern Anal. Mach. Intelligence,* vol. 23, no. 6, pp. 643-660, 2001.

[17] Mehmet Korkmaz and Nihat Yilmaz, "Face Recognition by Using Back Propagation Artificial Neural Network and Windowing Method," *Journal of Image and Graphics*, vol. 4, no.1, pp. 15-19, June 2016.

[18] Thangairulappan Kathirvalavakumar and Jeyasingh Jebakumari Beulah Vasanthi, "Face Recognition Based on Wavelet Packet Coefficients and Radial Basis Function Neural Networks," *Journal of Intelligent Learning Systems and Applications*, vol. 5 no. 2, pp. 115-122, 2013.

[19] Alex Pappachen James, Irina Fedorova, Timur Ibrayev and Dhireesha Kudithipudi, "HTM Spatial Pooler With Memristor Crossbar Circuits for Sparse Biometric Recognition,"*IEEE Transactions on Biomedical Circuits and Systems,* vol. 11, issue 3, pp. 640-651, June 2017.

[20] Robert Yao Aaronson, Wu Chen and Ben-Bright Benuwa, "Robust Face Detection using Convolutional Neural Network," *International Journal of Computer Applications*, vol. 170, no. 6, pp. 14-20, July 2017.

[21] Veena Mayya, Radhika M. Pai and Manohara Pai M.M ,"Automatic Facial Expression Recognition Using DCNN," *International Conference on Advances in Computing &Communications*, vol. 93,pp.453-461,2016.

[22] Bingrong Xu, Qingshan Liu and Tingwen Huang, "A Discrete-Time Projection Neural Network for Sparse Signal Reconstruction With Application to Face Recognition," *IEEE Transaction on Neural Networks and Learning Systems*, vol. PP, issue 99, pp. 1-12, June 2018.

[23] Qingfeng Liu and Chengjun Liu, "A Novel Locally Linear KNN Method With Applications to Visual Recognition," *IEEE Transaction on Neural Networks and Learning Systems*, vol. 28, issue 9, pp. 2010-2021, Sept. 2017.