International Journal of Engineering Innovation & Research
Volume 1, Issue 1, ISSN : 2277 – 5668

Performance Analysis of Face Recognition by Principal Component Analysis and Feed Forward Neural Network

S.N. Kakarwal , R.R. Deshmukh

***Abstract* — In face recognition, it is important to select the invariant facial features especially faces with various poses and expression change. This paper presents novel technique for recognizing faces viz. PCA + FFNN. The experiment is performed over FERET faces. This technique gives results with considerable accuracy.**

***Key Words* — Biometric, PCA and FFNN, Pattern Matching, FERET.**

I. Introduction

Face recognition from still images and video sequence has been an active research area due to both its scientific challenges and wide range of potential applications such as biometric identity authentication, human-computer interaction, and video surveillance. Within the past two decades, numerous face recognition algorithms have been proposed as reviewed in the literature survey. Even though we human beings can detect and identify faces in a cluttered scene with little effort, building an automated system that accomplishes such objective is very challenging. The challenges mainly come from the large variations in the visual stimulus due to illumination conditions, viewing directions, facial expressions, aging, and disguises such as facial hair, glasses, or cosmetics [1]. Face Recognition focuses on recognizing the identity of a person from a database of known individuals. Face Recognition will find countless unobtrusive applications such as airport security and access control, building surveillance and monitoring Human-Computer Intelligent interaction and perceptual interfaces and Smart Environments at home, office and cars [2].

Within the last decade, face recognition (FR) has found a wide range of applications, from identity authentication, access control, and face-based video indexing/browsing, to human-computer interaction.

Two issues are central to all these algorithms: 1) feature selection for face representation and 2) classification of a new face image based on the chosen feature representation. This work focuses on the issue of feature selection. Among various solutions to the problem, the most successful are those appearance-based approaches, which generally operate directly on images or appearances of face objects and process the images as two-dimensional (2-D) holistic patterns, to avoid difficulties associated with three-dimensional (3-D) modeling, and shape or landmark detection [3]. The initial idea and early work of this research have been published in part as conference papers in [4], [5] and [6].

A recognition process involves a suitable representation, which should make the subsequent processing not only computationally feasible but also robust to certain variations in images. One method of face representation attempts to capture and define the face as a whole and exploit the statistical regularities of pixel intensity variations [7].

The remaining part of this paper is organized as follows. Section II extends to the pattern matching which also introduces and discusses the Principal Component Analysis and FFNN in detail. In Section III, extensive experiments on FERET are conducted to evaluate the performance of the proposed method on face recognition. Finally, conclusions are drawn in Section IV with some discussions.

II. Pattern Matching

A. Pattern Recognition Methods

During the past 30 years, pattern recognition has had a considerable growth. Applications of pattern recognition now include: character recognition; target detection; medical diagnosis; biomedical signal and image analysis; remote sensing; identification of human faces and of fingerprints; machine part recognition; automatic inspection; and many others.

Traditionally, Pattern recognition methods are grouped into two categories: structural methods and feature space methods. Structural methods are useful in situation where the different classes of entity can be distinguished from each other by structural information, e.g. in character recognition different letters of the alphabet are structurally different from each other. The earliest-developed structural methods were the syntactic methods, based on using formal grammars to describe the structure of an entity [8].

The traditional approach to feature-space pattern recognition is the statistical approach, where the boundaries between the regions representing pattern classes in feature space are found by statistical inference based on a design set of sample patterns of known class membership [8]. Feature-space methods are useful in situations where the distinction between different pattern classes is readily expressible in terms of numerical measurements of this kind. The traditional goal of feature extraction is to characterize the object to be recognized by measurements whose values are very similar for objects in the same category, and very different for objects in different categories. This leads to the idea of seeking distinguishing features that are invariant to irrelevant transformations of the input. The task of the classifier component proper of a full system is to use the feature vector provided by the feature extractor to assign the object to a category [9]. Image classification is implemented by computing the similarity score between a target discriminating feature vector and a query discriminating feature vector [10].

B. Wavelet Transform

Wavelets decompose complex signals into sums of basis functions – in this respect they are similar to other discrete image transforms. However, wavelets are local in both frequency and time and are able to analyze data at different scales or resolutions much better than simple sine and cosine can [11].

Wavelets are an extension of Fourier analysis. The goal is to turn the information of a signal into numbers – coefficients – that can be manipulated, stored, transmitted, analyzed, or used to reconstruct the original signal. There are not only two big classes of wavelet transforms - continuous and discrete - but discrete transforms can be redundant, orthogonal, or biorthogonal. Each category contains innumerable possibilities, Daubechies wavelets alone constituting a very big class [12].

DWT for an image as a 2-D signal can be derived from 1-D DWT. The easiest way for obtaining scaling and wavelet functions for two dimensions is by multiplying two 1-D functions. The scaling functions for 2-D DWT can be obtained by multiplying two 1-D scaling functions: $\phi(x,y)=\phi(x)\phi(y)$. Wavelet functions for 2-D DWT can be obtained by multiplying two wavelet functions. For the 2-D case, there exist three wavelet functions that scan details in horizontal $\psi(\mathrm{I})(x,y)=\varphi(x)\psi(y)$, vertical $\psi(\mathrm{II})(x,y)=\psi(x)\phi(y)$, and diagonal directions: $\psi(\mathrm{III})(x,y)=\psi(x)\psi(y)$ [13]. In this paper wavelet transform is used to compress the images.

C. Principal Component Analysis

In some situations, the dimension of the input layer is large, but the components of the vectors are highly correlated (redundant). It is useful in this situation to reduce dimension of the input vectors. An effective procedure for performing this operation is principal component analysis. This technique has three effects: it orthogonalizes the components of the input vectors (so that they are uncorrelated with each other); it orders the resulting orthogonal components (principal components) so that those with the largest variation come first; and it eliminates those components that contribute the least to the variation in the data set.We have used prepca (Matlab function) to perform principal component analysis. Note that we first normalize the input vectors, using prestd (Matlab function), so that they have zero mean and unity variance.

In this paper we have used singular value decomposition to compute the principal components. The input vectors are multiplied by a matrix whose rows consist of the eigenvectors of the input covariance matrix. This produces transformed input vectors whose components are uncorrelated and ordered according to the magnitude of their variance. Those components which contribute only a small amount to the total variance in the data set are eliminated. It is assumed that the input data set has already been normalized so that it has a zero mean. The function prestd can be used to normalize the data [14].

The current state of feature extraction is characterized by a family of subspace methods originated by eigenface, the core of which is principal component analysis (PCA) [15].

The features are extracted from faces by PCA.

We can use SVD to perform PCA. We decompose X using SVD,

i.e. $X = UTV^T$ and that we can write the covariance matrix as

$C = 1/n\, XX^T = 1/n\, UT^2U^T$

In this case U is a *m x n* matrix. Following from the fact that SVD routine order the singular values in descending order we know that, if n <*m*, the first *n* columns in U corresponds to the sorted eigenvalues of C and if *m*≥¸ *n*, the first *m* corresponds to the sorted non-zero eigenvalues of C. The transformed data can thus be written as

$$Y = \tilde{U}^T X = \tilde{U}^T U^T V^T$$

where $\tilde{U}^T X$ is a simple *m x n* matrix which is one on the diagonal and zero everywhere else. To conclude, we can write the transformed data in terms of the SVD decomposition of X [16].

D. Artificial Neural Network

In recent years, there has been an increase in the use of evolutionary approaches in the training of artificial neural networks (ANNs). While evolutionary techniques for neural networks have shown to provide superior performance over conventional training approaches, the simultaneous optimization of network performance and architecture will almost always result in a slow training process due to the added algorithmic complexity [17].

E. Feed Forward Network

Feed forward networks may have a single layer of weights where the inputs are directly connected to the output, or multiple layers with intervening sets of hidden units. Neural networks use hidden units to create internal representations of the input patterns [18].

A Feed forward artificial neural network consists of layers of processing units, each layer feeding input to the next layer in a Feed forward manner through a set of connection weights or strengths. The weights are adjusted using the back propagation learning law. The patterns have to be applied for several training cycles to obtain the output error to an acceptable low value.

The back propagation learning involves propagation of the error backwards from the input training pattern, is determined by computing the outputs of units for each hidden layer in the forward pass of the input data. The error in the output is propagated backwards only to determine the weight updates [19]. FFNN is a multilayer Neural Network, which uses back propagation for learning [20].

International Journal of Engineering Innovation & Research
Volume 1, Issue 1, ISSN : 2277 – 5668

As in most ANN applications, the number of nodes in the hidden layer has a direct effect on the quality of the solution. ANNs are first trained with a relatively small value for hidden nodes, which is later increased if the error is not reduced to acceptable levels. Large values for hidden nodes are avoided since they significantly increase computation time [20].

The Back propagation neural network is also called as generalized delta rule. The application of generalized delta rule at any iterative step involves two basic phases. In the first phase, a training vector is presented to the network and is allowed to propagate through the layers to compute output for each node. The output of the nodes in the output layers is then compared against their desired responses to generate error term. The second phase involves a backward pass through a network during which the appropriate error signal is passed to each node and the corresponding weight changes are made. Common practice is to track network error, as well as errors associated with individual patterns. In a successful training session, the network error decreases with the number of iterations and the procedure converges to a stable set of weights that exhibit only small fluctuations with additional training. The approach followed to establish whether a pattern has been classified correctly during training is to determine whether the response of the node in the output layer associated with the pattern class from which the pattern was obtained is high, while all the other nodes have outputs that are low [21].

III. Experiments

A. Facial Database

There are many facial databases available for evaluating face recognition algorithms. The FERET facial database consists of 13539 facial images corresponding to 1,565 subjects. Since images are acquired during different photo sessions, the illumination conditions and the size of the face may vary. The diversity of the FERET database is across gender, ethnicity, and age. The images are acquired without any restrictions imposed on facial expression and with at least two frontal images shot at different times during the same photo session. The FERET database has become the de facto standard for evaluating face recognition technologies [2]. The FERET dataset used in our experiments includes 250 face images corresponding to 20 different subjects. In order to extract the facial region, the images are normalized. All images are gray-scale images.

B. Steps used in Face Recognition

1. Read the Images
2. Convert the color images into gray scale images
3. Perform pre-processing
3. Apply DWT to compress the images
4. Compute Principal Component Analysis of image feature space
5. Classify the images by Feed forward neural network with different values of number of epochs
6. Analyse the performance by computing FAR and FRR at different values of threshold with different values of dimensionality reduction

C. Performance Evaluation

The accuracy of biometric-like identity authentication is due to the genuine and imposter distribution of matching. The overall accuracy can be illustrated by Receiver Operation Characteristics (ROC) curve, which shows the dependence of False Reject Rate (FRR) on False Accept Rate (FAR) at all thresholds. When the parameter changes, FAR and FRR may yield the same value, which is called Equal Error Rate (EER). It is a very important indicator to evaluate the accuracy of the biometric system, as well as binding of biometric and user data [22].

A typical biometric verification system commits two types of errors: *false match* and false *non-match*. Note that these two types of errors are also often denoted as *false acceptance* and *false rejection*; a distinction has to be made between positive and negative recognition; in positive recognition systems (e.g., an access control system) a false match determines the false acceptance of an impostor, whereas a false non-match causes the false rejection of a genuine user. On the other hand, in a negative recognition application (e.g., preventing users from obtaining welfare benefits under false identities), a false match results in rejecting a genuine request, whereas a false non-match results in falsely accepting an impostor attempt. The notation “false match/false non-match” is not application dependent and therefore, in principle, is preferable to “false acceptance/false rejection.” However, the use of false acceptance rate (FAR) and false rejection rate (FRR) is more popular and largely used in the commercial environment [23].

Traditional methods of evaluation focus on collective error statistics such as EERs and ROC curves. These statistics are useful for evaluating systems as a whole. *Equal-Error Rate* (EER) denotes the error rate at the threshold t for which false match rate and false non-match rate are identical: $FAR(t) = FRR(t)$ [24].

Finally, Table I represents the genuine and impostor score w. r. t. different values of threshold, percentage of dimensionality reduction and number of epochs. FAR and FRR values for all persons with different threshold values. The FRR and FAR for number ofparticipants (N) are calculated as specified in Eq. (1) and in equation Eq. (2) [25]:

$$FRR = \frac{1}{N}\sum_{n=1}^{N} FRR(n) \qquad \ldots(1)$$

$$FAR = \frac{1}{N}\sum_{n=1}^{N} FAR(n) \qquad \ldots(2)$$

Conclusions

This paper investigates the feasibility and effectiveness of face recognition with PCA and FFNN. A 2D Discrete Wavelet Transform is used to compress the faces. Face recognition based on PCA is performed by back propagation and thresholding. Experimental results on FERET database demonstrate that the proposed methodology outperforms in recognition. Experiments have been conducted on various face conditions, including different angles, expressions etc.

The face recognition system is trained with different values of number of epochs such as 17000, 17100 and 17200, and different values of dimensionality reduction such as 0.2% and 0.3%. Among all these values best results are obtained when number of epochs=17000 and percentage of dimensionality reduction of 0.2%

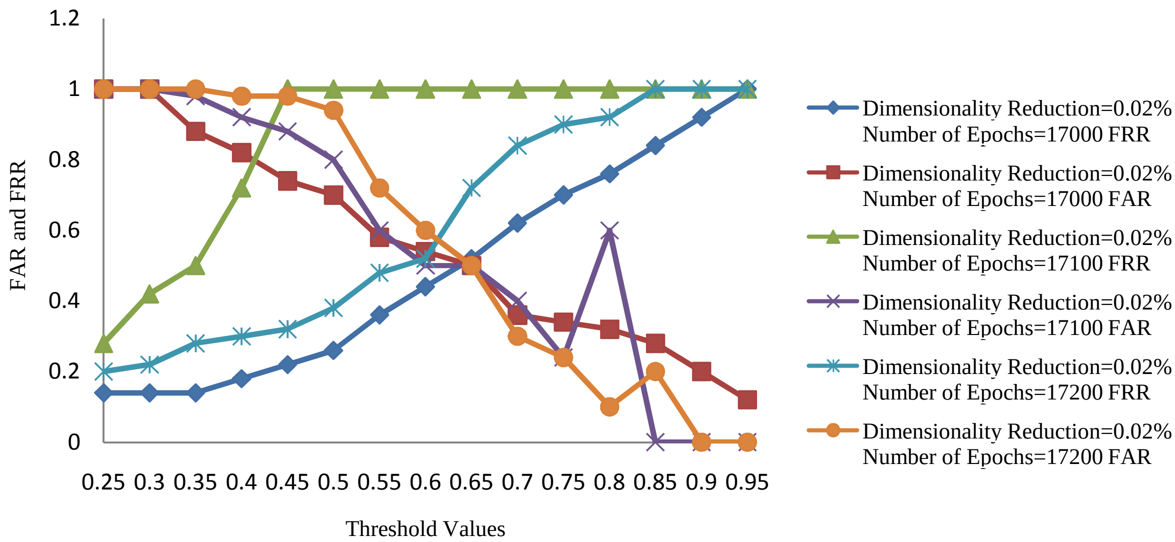
There is a trade-off between FAR and FRR values. When number of epochs is 17000, the optimum thresholds value i.e. the value where FAR=FRR, is 0.65. For threshold value 0.65, the value of FRR is 0.52 and the value of FAR is 0.5 as shown in Graph1. The Graph2 shows the plot of FRR and FAR when percentage of dimensionality reduction is 0.3%.

ACKNOWLEDGEMENTS

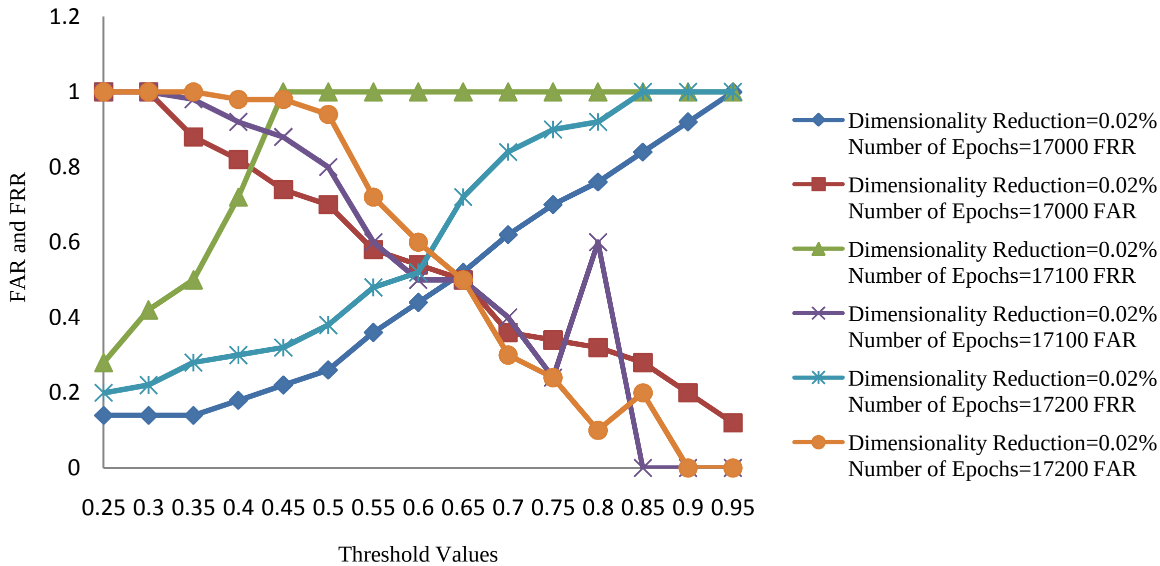
The authors would like to thank the anonymous reviewers for their critical and constructive comments and suggestions.

Table1: Genuine Score and Impostor Score

Threshold Value	Dimensionality Reduction											
	0.2%						0.3%					
	Number of Epochs						Number of Epochs					
	17000		17100		17200		17000		17100		17200	
	FRR	FAR	FRR	FAR	FRR	FAR	FRR	FAR	FRR	FAR	FRR	FAR
0.25	0.14	1	0.28	1	0.2	1	0.14	1	0.28	1	0.54	0.98
0.3	0.14	1	0.42	1	0.22	1	0.16	1	0.36	1	0.7	0.98
0.35	0.14	0.88	0.5	0.98	0.28	1	0.22	1	0.36	1	0.8	0.9
0.4	0.18	0.82	0.72	0.92	0.3	0.98	0.26	0.98	0.38	0.98	0.88	0.78
0.45	0.22	0.74	1	0.88	0.32	0.98	0.3	0.86	0.38	0.96	0.9	0.46
0.5	0.26	0.7	1	0.8	0.38	0.94	0.38	0.84	0.5	0.9	0.96	0.36
0.55	0.36	0.58	1	0.6	0.48	0.72	0.38	0.72	0.56	0.66	0.98	0.3
0.6	0.44	0.54	1	0.5	0.52	0.6	0.42	0.54	0.6	0.48	1	0.26
0.65	0.52	0.5	1	0.5	0.72	0.5	0.42	0.14	0.66	0.46	1	0.12
0.7	0.62	0.36	1	0.4	0.84	0.3	0.54	0	0.68	0.44	1	0.8
0.75	0.7	0.34	1	0.24	0.9	0.24	0.66	0	0.72	0.44	1	0.6
0.8	0.76	0.32	1	0.6	0.92	0.1	0.8	0	0.78	0.38	1	0.4
0.85	0.84	0.28	1	0	1	0.2	0.82	0	0.92	0.36	1	0.4
0.9	0.92	0.2	1	0	1	0	0.94	0	0.92	0.26	1	0.2
0.95	1	0.12	1	0	1	0	0.96	0	0.98	0.14	1	0



Graph1: Genuine and Impostor Score with dimensionality reduction = 0.2%



Graph2: Genuine and Impostor Score with dimensionality reduction =0.3%

REFERENCES

[1] Yu Su, Shiguang Shan, Xilin Chen, and Wen Gao," Hierarchical Ensemble of Global and local Classifiers for Face Recognition", IEEE Transactions On Image Processing,Vol. 18, No. 8, August 2009, pp 1885 – 1896

[2] Chengjun Liu, Harry Wechsler, "Independent Componenet Analysis of Gabor Features for Face Recognition", IEEE Transactions on Neural Networks, Vol. 14, July 2003, pp 919-928

[3] Juwei Lu, Konstantinos N. Plataniotis, Anastasios N. Venetsanopoulos, "Face Recognition Using Kernel Direct Discriminant Analysis Algorithms, IEEE Transactions On Neural Networks", Vol. 14, No. 1, January 2003, pp 117-126

[4] S.N. Kakarwal, S.D. Sapkal, P.J.Ahire, Dr. D.S. Bormane," Analysis of Facial Image Classification using Discrete Wavelet Transform", Proc. International Conference ICSCI, 2007, pp 700-705

[5] S.N. Kakarwal, Mahananda Malkauthekar, Shubhangi Sapkal, and Dr. Ratnadeep Deshmukh,"Face Recognition using Fourier Descriptor and FFNN", Proc. IEEE International Advance Computing Conference, 2009, pp 2740-2742

[6] S.N. Kakarwal, Dr. R.R. Deshmukh, "Wavelet Transform based Feature Extraction for Face Recognition", pp 100-104, IJCSA Issue-I June 2010, ISSN 0974-0767

[7] Bai-Ling Zhang, Haihong Zhang and Shuzhi Sam Je, "Face Recognition by Applying Subband Representation and Kernel Associative Memory", IEEE Transaction on Neural Networks, Vol. 15, Jan 2004, pp 166-177

[8] Daisheng Luo, "Pattern Recognition and Image Processing",Horwood Publishing Limited 1998, pp 2-3

[9] Richard O. Duda, Peter E. Hart, David G. Stork,"Pattern Classification", John Wiley 2001, pp 11-12

[10] Chengjun Liu, "Learning the Uncorrelated, Independent, and Discriminating Color Spaces for Face Recognition", IEEE Transactions On Information Forensics And Security, Vol. 3, No. 2, June 2008, pp 213-222

[11] Milan Sonka, "Image Processing,Analysis, and Machine Vision",Thomson Learning 2001, pp 617

[12] Barbara, "The World According to Wavelets", pp 30

[13] Sonja Grgic, Mislav Grgic, Branka Zovko-Cihlar,"Performance Analysis of Image Compression Using Wavelets", IEEE Transactions on Industrial Electronics, Vol. 48, No. 3, June 2001, pp 682-695

[14] Howarad Demuth, Mark Beale, Neural Network toolbox User's Guide Version4, pp 5-63

[15] Yanwei Pang, Yuan Yuan, and Xuelong Li, Gabor-Based Region Covariance Matrices for Face Recognition, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 18, No. 7,pp. 989-993, July 2008

[16] Rasmus Elsborg Madsen, Lars Kai Hansen and Ole Winther, Singular Value Decomposition and Principal Component Analysis, pp 04

[17] Chi-Keong Goh, Eu-Jin Teoh, and Kay Chen Tan, Hybrid Multiobjective Evolutionary Design for Artificial Neural Networks, IEEE Transactions on Neural Networks, Vol. 19, no. 9, September 2008

[18] S.N. Sivanandanam, S. Sumathi, S. N. Deepa, Introduction to Neural Networks using MATLAB 6.0, Tat McGrawHill Publishing Company Limited, pp.20

[19] B. Yegnanarayana, Artificial Neural Networks, PHI, pp 231

[20] Kevin Stanley McFall and James Robert Mahan, Artificial Neural Network Method for Solution of Boundary Value Problems with Exact Satisfaction of Arbitrary Boundary Conditions, IEEE Transactions on Neural Networks, Vol. 20, No. 8, pp. 1221-1233, August 2009

[21] Rafel C. Gonzalez, Richard E. Woods, Digital Image Processing, pp. 896

[22] Stan Z. Li, Anil K. Jain, Encyclopedia of Biometrics, Springer, pp 75

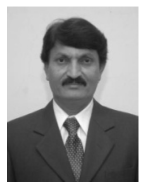
[23] Davide Maltoni, Dario Maio, Anil K. Jain, Salil Prabhakar: Handbook of Fingerprint Recognition(Springer), pp 3

[24] Neil Yager and Ted Dunstone: The Biometric Menagerie, IEEE Transactions On Pattern Analysis And Machine Intelligence,VOL.32,NO. 2, Feb2010, pp 220-230

[25] Website: www.bromba.com/faq/biofaqe

AUTHOR'S PROFILE

S.N. Kakarwal, received M.E. and B.E. degree in Computer Science and Engineering. She is currently working toward the Ph.D. degree under guidance of Dr. R.R. Deshmukh at Dr. Babasaheb Ambedkar Marathwada Univeristy, Aurangabad. Presently working as an Associate Professor in Department of Computer Science and Engineering, P.E.S. College of Engineering, Aurangabad, MS-India. She is a member of Computer Society of India. She was organizing committees of 'Tectrix2K7' and 'Pulses 2011'. Her research interests include Image Processing, Pattern Recognition and Artificial Neural Network. In these areas, she has published 25 research papers in leading Journals, National and International conferences proceedings. She has bagged 3 Best Paper Award.



Dr. R. R. Deshmukh, M.E. (CSE), M.Sc. (CSE) Ph.D. FIETE, Presently working as an Associate Professor in Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, MS-India. He is a fellow of IETE, Life member of ISCA, CSI, ISTE, ACEEE, IAEng, CSTA, IDES and a senior member of IEEE. He is a Member of Management Council of Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, MS-India. He is editor of four books and published more than 40 research papers in reputed Journals, National and international conferences. He is reviewer and editor of several journals at national & international level. He has organized several workshops and conferences. He is nominated as a subject expert on various academic & professional bodies at national level government bodies. He is working as Chairman, Ad-hoc Board of Computer Science & IT, Faculty member for Engineering, Science & Management Faculty & Member of various committees at university level. He has two research project projects from UGC and received grants more than 10 Lakhs. His areas of specialization are Data Mining, Data Warehousing, Image Processing, Pattern Recognition, Artificial Intelligence, Computational Auditory Scene Analysis (CASA), Neural Networks etc. He won First prize in Inter University State Level Research Festival "AVISHKAR - 2009" under H. L. F. A. category at Teacher level & for the Team Management.