International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 17 (2015) pp 38583-38588

A novel Facial Expression Classification System using Emotional Back Propagation Artificial Neural Network using Genetic Algorithm

Vedantham Ramachandran
Research Scholar,
Dept. of CSE,
Acharya Nagarjuna University,
Guntur, AP, India

Dr. E.Srinivasa Reddy
Professor & Dean,
Acharya Nagarjuna University,
Guntur, AP, India

B. Sivaiah
Assistant Professor,
Vignan's University,
Guntur, AP, India

Abstract—The research in this paper supports potential applications in a wide range of areas, including human to computer interaction (HCI), disease diagnosis, transportation security, and crime prevention etc. FER techniques help to develop emotionally intelligent HCI systems which can sense and respond appropriately to user's emotional feedbacks in a more natural, friendly, and trustworthy way. In hospitals, patients' ability of expressing or discriminating some emotions can help doctors make a correct, accurate, and quick diagnosis of certain diseases.In this paper anovel classification system for facial expressions with back propagation artificial neural network that uses Genetic algorithm for pre-processing in feature extraction of images is proposed. This technique is capable of processing images extremely rapidly while achieving high detection rates for facial expressions. This entire system deploys two important contributions that are optimally mixed to make the entire system swifter. A set of experiments in the domain of facial expression detection were conducted. The system yields facial expression classification performance comparable to the previous systems that are implemented on a standard desktop.

Keywords—Facial Emotion recognition, Back Propagation Neural Network, Genetic Algorithm

I. INTRODUCTION

Facial expression is an important channel of human social communication. Facialexpressions can help humans perceive useful information, make correct decisions,and give instant responses during social communications. Studies havedemonstrated that facial expressions contribute for 55% to the effect of the spokenmessage. Facial expression recognition (FER) aims to perceive and understandemotional states of humans from the face. It is a main component of the emergingaffective computing,which focuses on perceiving, interpreting and imitating theuser's affective states, then responding appropriately to the affective signals.Building automatic and robust FER systems also plays a critical role for the nextgenerationcomputing - human computing, where the user interfaces or modelsare human-centred, and they are capable of sensing and responding appropriately touser's affective feedbacks in a more natural, efficacious, persuasive, friendly, andtrustworthy way[1].

Development of a robust FER system is still a challenging issue, largelybecause of various unpredictable facial variations and complicated exteriorenvironmental conditions. These variations make it difficult to pre-locate facialregions and perform robust and accurate feature extraction. Many efforts havebeen made to overcoming these variations, especially in pose and illuminationcondition. Although these studies have achieved promising results, to be best of ourknowledge, no approach has been reported on handling face localization errors (e.g.changes in face scale and location), and relatively little attention on overcomingfacial occlusions. The significant impact of these two types of variations on FERperformance has been specifically highlighted in recent works [1].

Recently, the facial expression recognition technology attracts more and more attention with people's growing interesting in expression information. Facial expression carries crucial information about the mental, emotional and even physical states of the conversation. Facial expression recognition has practical significance; it has very broad application prospects, such asuser-friendly interface between man and machine, humanisticdesign of goods, and emotional robot etc. With facialexpression recognition systems, the computer will be able toassess the human expressions depending on their effectivestate in the same way that human's senses do. Theintelligent computers will be able to understand, interpret andrespond to human intentions, emotions and moods [2].

Real-world FER systems normally require a rigid pre-processing of the face toovercome facial and environmental variations, and a careful selection of features toextract subtle information of naturally expressed emotions. The investigation oftechniques to handle face localization errors and facial occlusions can relieve therequirements in face pre-processing, while comparisons of different types of featureshelp to extract more useful information. In addition, combination of texture andgeometric features, and selection of the most discriminative feature subset can alsobe used to improve the recognition performance. Evaluations of FER algorithms onreal-world data can provide useful insights into differences between standard imagesand real-world images, and facilitate developing FER systems for real-worldapplication. However, little attention has been paid to addressing the above issues.

The research in this paper supports potential applications in a wide range ofareas, including human to computer interaction (HCI), disease diagnosis,transportation security, and crime prevention etc. FER techniques help to developemotionally intelligent HCI systems which can sense and respond appropriately touser's emotional feedbacks in a more natural, friendly, and trustworthy way. Inhospitals, patients' ability of expressing or discriminating some emotions can helpdoctors make a correct, accurate, and quick diagnosis of certain diseases. FERtechniques have been applied to improve health care, differentiate schizophreniafrom other problems, ascertain the level of Parkinson's disease. In publictransportation, monitoring drivers' fatigue level can help to issue a warning beforepossible accidents. At the police station, discriminating between posed andspontaneous facial expressions can be used as a complementary tool to spot a liar. Atthe airport, detecting facial expression changes can help to identify those people withbad intentions and prevent the possible crimes [1].

Figure 1 shows the framework of a general FER research and development system.It contains six main steps: face pre-processing (face detection, tracking, andnormalization), feature extraction, feature selection, emotion classification, emotionrepresentation, and performance evaluation. The input data can be static images orvideo sequences. The classified facial expressions are then represented in differentmethods and the performance is evaluated using different measurements.

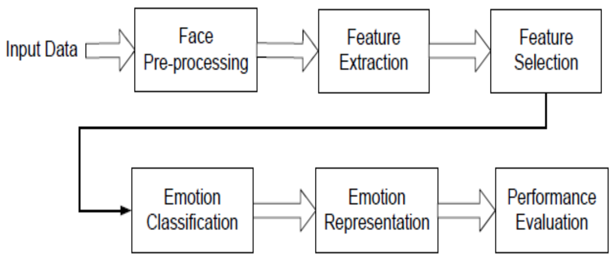


Figure 1An Expression Recognition system

II. PROPOSED FACIAL RECOGNITION SYSTEM

The proposed architecture in this work contains the following stages: preprocessing of input images, feature extraction, training, classification, and database.This work proposes a new solution tothe facial expression recognition problem, describing afacial recognition system that can be used in application ofHuman computer interface.Pre-processing of input images includes, face detection and cropping. Feature extraction is the process of deriving unique features from the data and can be accomplished by specific algorithms like Feature averaging, principal component analysis etc.Training of neural network will be doneby giving the extracted features as input to the neural network with specified network parameters.

Classification will be done by the neural network according to the specified targets in the network. The following figure shows the proposed architecture of the Facial recognition system.Pre-processing is the next stage after entering the data intothe facial expression recognition system. The important datathat is needed for most facial expression recognition methodsis face position. In pre-processing module images are resized from 256 x 256 pixel value to 280 x 180 pixel values. TheSobel method has been used to identify the face edges.

Face images are taken from Cohn Kanade database of facial expressions. The original image contains time and camera model also. For better performance, face is detected and cropped and saved as separate image. The cropped image is then used to extract features. These features are given as input to the neural network and will be trained to gain knowledge.

Preprocessing:

The testing image will also be preprocessed and features will be extracted and input to the neural network. The classifier of the neural network will classify the expression of the input test image.

Figure 2 Face detection and cropping

In order to perform data reduction, the first step is to take the required data from an image. So the face is detected and cropped from original image as shown in Fig.2.

Genetic Algorithms

The Genetic Algorithms (GA´s) are characterized by a search technique inspired in the Evolutionist Theory by Darwin, uses some selection mechanism, where individuals that is the chromosomes more adapted of a population are the ones that have more survival chances, by getting used easily to changes that occur in its environment. This makes the algorithm strong and fast, being designate to a determined type of optimization, where the search space is too big and the conventional methods become inefficient. Another GA´s important characteristic is that they result a set of solutions and not only one solution.

a. An algorithm and the cycle of the GA are described in the following steps:
b. Start a population of N size, with chromosomes generated randomly.
c. Apply fitness to each chromosome of population.
d. Make new chromosomes through crossings of selected chromosomes of this population. Apply recombination and mutation in these chromosomes.
e. Eliminate members of old population, in order to have space to insert these new chromosomes, keeping the population with the same N chromosomes.
f. Apply fitness in these chromosomes and insert them in the population.
g. If the ideal solution will be found or, if the time (or generation number) depleted, return the chromosome with best fitness. Otherwise, come back to the step c.

The neural network will produce the knowledge database. In the process of testing, the test input image will be applied pattern averaging and the remaining features will be used to classify through the neural network classifier and using the knowledge database gained from training. The architecture is shown in the Fig.4.

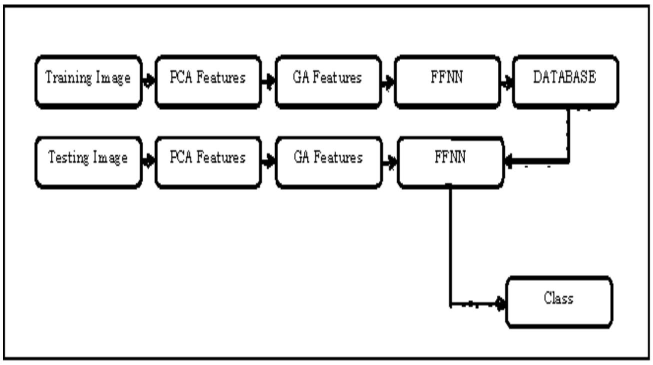


Figure 5 Architecture of FFNN classification with PCA features optimized by GA

GA Coefficients with Emotional Back Propagation Neural Network:

This architecture proposes the classification with neural network using GA Coefficients selected from PCA features of input images. The training images were taken and applied the PCA and then GA is used to select optimized features. The remaining features are input to the emotional back propagation neural network to train the network. The neural network will produce the knowledge database.

In the process of testing, the test input image will be applied pattern averaging and the remaining features will be used to classify through the neural network classifier and using the knowledge database gained from training. The architecture is shown in the Fig.

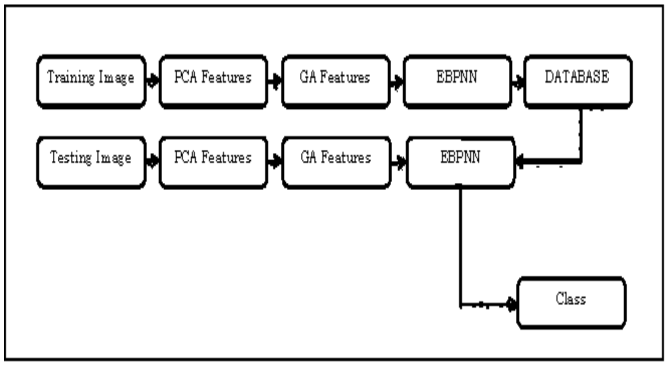


Figure 2 Proposed architecture of EBPNN classification with DCT Coefficients

The proposed architectures use the feed forward and emotional back propagation neural network architectures. The generalized architecture of the proposed system is shown in Fig.

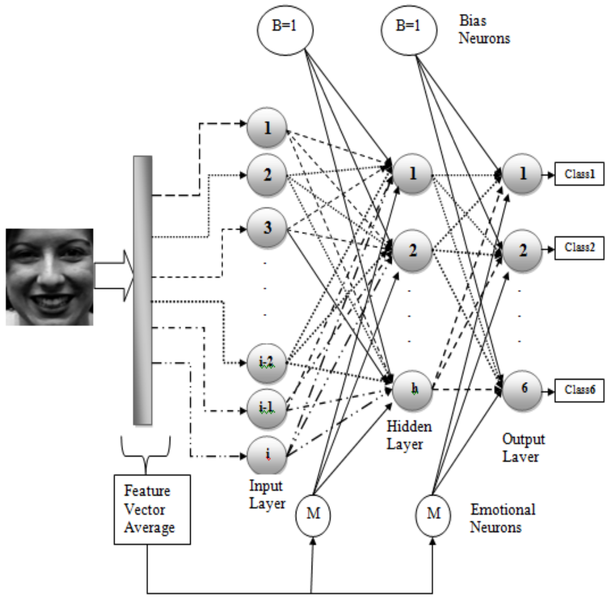


Figure 8 Generalized EmBP-based neural network

III. TESTS, RESULTS & CONCLUSIONS

The implementation of neural network consists of training and testing. The training and testing is performed on Cohn Kanade facial expression database. The database consists of 2000 images of 200 subjects. About 600 images were used in this work for the training and testing process. Sample images from the Cohn Kanade database are shown in Fig.4.

The performance of the system is measured by varying the number of images of each expression in training and testing. Following table shows the performance of the proposed method along with the other methods.

Table 1 Comparison of results on Cohn Kanade database

Training samples	Testing samples	Training Time		Recognition Rate	
		GA+FF NN	GA+BP NN	GA+FF NN	GA+BP NN
8	2	160.85	152.30	100	100
7	3	144.29	202.39	100	100
6	4	69.50	84.43	96	98
5	5	25.145	33.69	87	97
4	6	8.86	13.08	85	96
3	7	3.37	4.48	83	95
2	8	2.88	4.17	80	95

The recognition performance increases as the number of training samples increases. The lower the number of training samples the lesser the recognition rate. It is found that the DCT with emotional back propagation neural network is yielding the better results even the training samples are less. The performance plot was shown against various algorithms, number of training images and their performances.

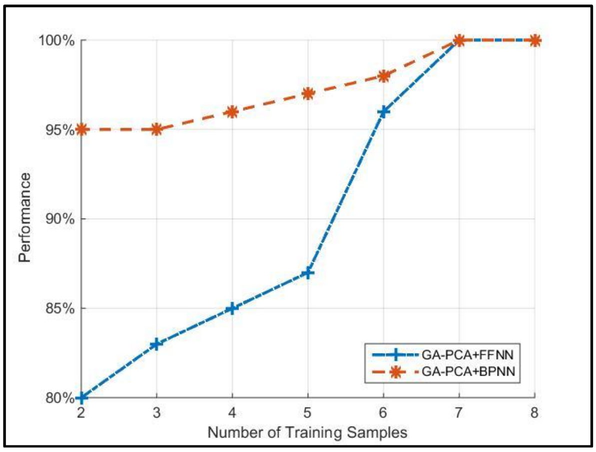


Figure 9 Plot against FFNN & BPNN Recognition Rate

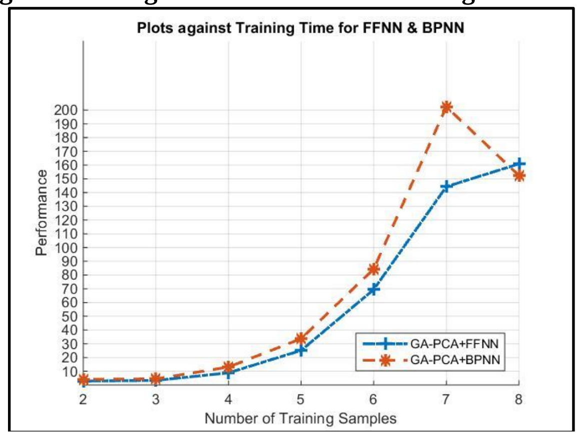


Figure 10 Plot against FFNN & BPNNTraining Times

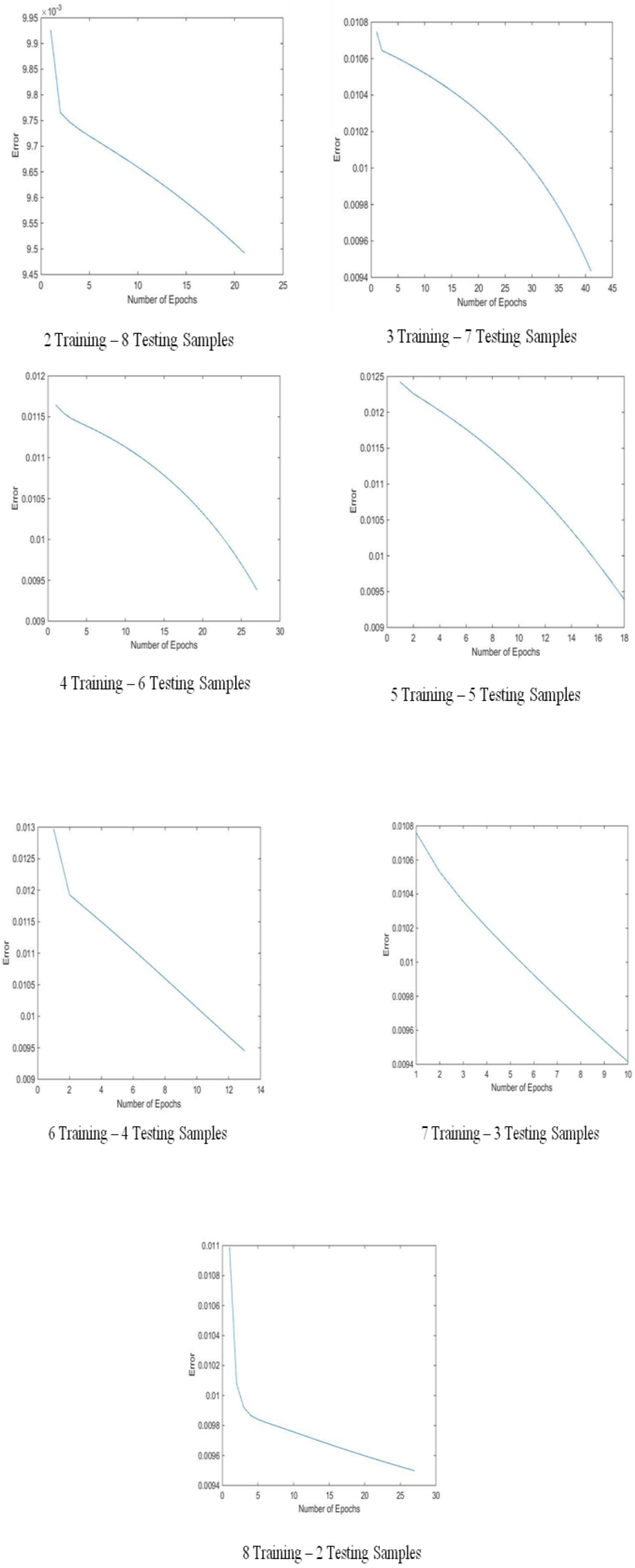
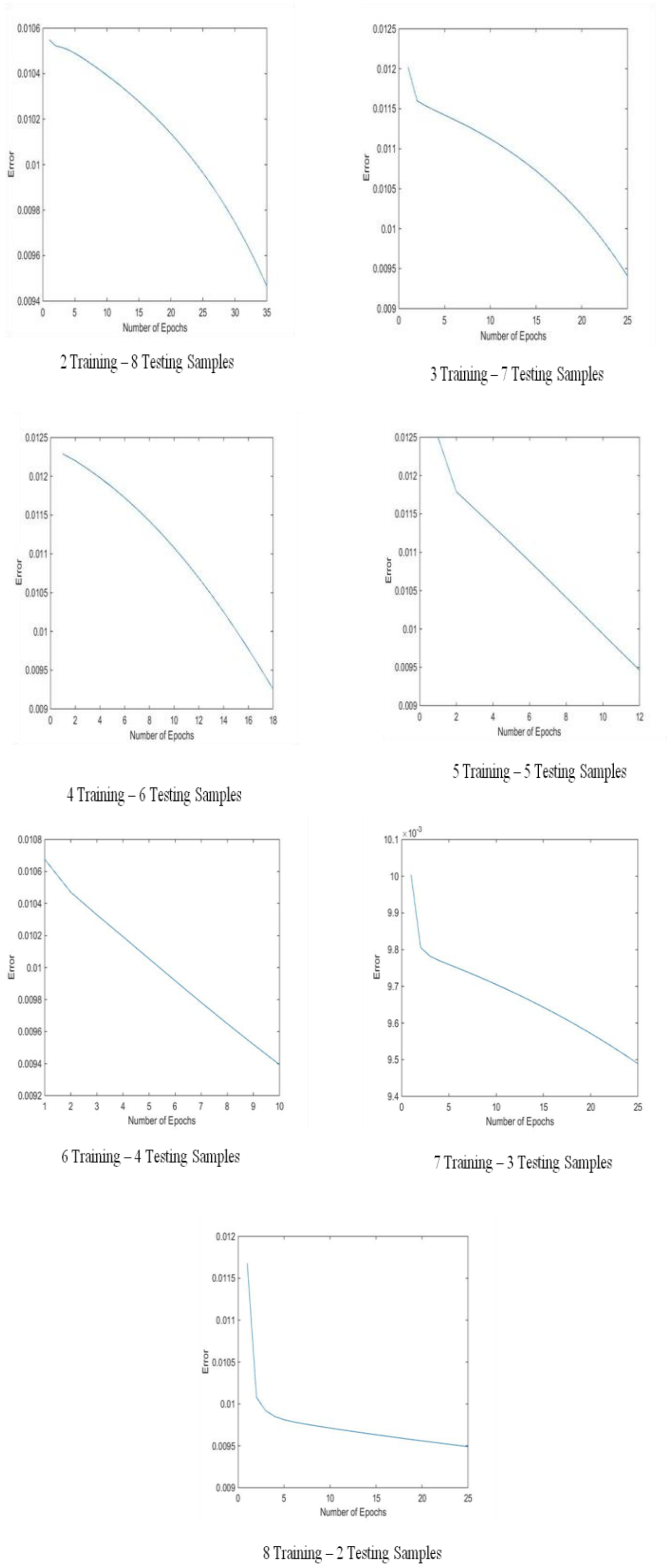


Fig 11. Error Minimization Plots for EBPNN



2 Training – 8 Testing Samples

3 Training – 7 Testing Samples

4 Training – 6 Testing Samples

5 Training – 5 Testing Samples

6 Training – 4 Testing Samples

7 Training – 3 Testing Samples

8 Training – 2 Testing Samples

Fig 12. Error Minimization Plots for FFNN

The confusion matrix is created for each of the test. The test is performed on five subjects.

Found / Actual	Surprise	Fear	Happy	Sad	Anger	Disgust
Surprise	100	0	0	0	0	0
Fear	0	100	0	0	0	0
Happy	0	0	60	0	0	40
Sad	0	0	0	100	0	0
Anger	0	0	0	60	40	0
Disgust	0	0	0	0	0	100

Figure 13GA-PCA+BPNN Confusion Matrix

The confusion matrix shows the percentage of correct classifications and mis-classifications also. Diagonal elements show the correct classification results.

All other elements are the misclassifications. The test and training results of various facial emotional classification methods is shown in Fig. Experimental results show that the proposed architecture improves the performance of the facial expressions. Based on the results we can conclude that the proposed emotional back propagation neural network with DCT is best in both cases of minimization of training time of neural network and performance as well. Since the emotional parameters were introduced, the training time for the single iteration may be little more but the overall training time is reduced in achieving the minimization of error.

The performance and training time of the neural network depends on the parameters selected like learning coefficient and momentum factor. The number of hidden neurons is also affecting the performance of the neural network. Experiments were carried out by altering the learning coefficient and number of hidden neurons and the types of sigmoid functions.

The optimal value for learning rate is 0.02, which produces the best performance for facial expression recognition. The number of hidden neurons is same as the number of input neurons. Sigmoid action function is used in both hidden layer and output layer for activating the neurons. In the classification part of the emotional back propagation neural network, the time very less when compared to other neural networks.

The work can be extended to clustering techniques like segmentation for the lower training times and higher performance. Since the training data is still images, there is more dependency on the image data like lighting, illumination conditions, poses of the faces, variations in expression and gender of the person also.

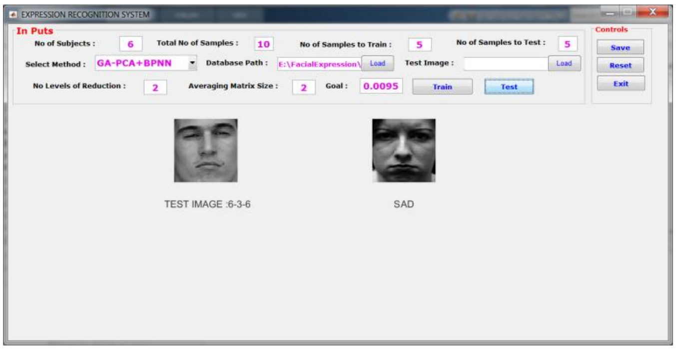


Fig 14. GUI Result for the Developed System

Conclusions and Future Work

We present an approach for facial expression estimation that combines state-of-the-art techniques for model-based image interpretation and sequence labelling. Learned objective functions ensure robust model-fitting to extract accurate model parameters. The classification algorithm is explicitly designed to consider sequences of data and therefore considers the temporal dynamics of facial expressions. Future work aims at presenting the classifier training data that is obtained from various publicly available databases to reflect a broader. Variety of facial expressions. Furthermore, our approach will be tuned towards applicability in real-time. It is planned tocreate a working demonstrator from this approach.

References

1. Adnan Khashman , "A modified back propagation learning algorithm with added emotional coefficients", IEEE transactions on neural networks, vol. 19, no. 11, November 2008
2. D. Rumelhart, G. E. Hinton, and R. J.Williams, "Learning internal representations by error propagation," in Parallel Distributed Processing, D. Rumelhart and J. McClelland, Eds. Cambridge, MA: MIT Press, 1986, vol. 1.
3. P. M. Chavan, Manan C .Jadhav, Jinal B. Mashruwala, Aditi K. Nehete, Pooja A. Panjari, "Real Time Emotion Recognition through Facial Expressions for Desktop Devices", International Journal of Emerging Science and Engineering (IJESE) Volume-1, Issue-7, May 2013
4. Ajit P. Gosavi, S. R. Khot, "Facial Expression Recognition Using Principal Component Analysis", International Journal of Soft Computing and Engineering (IJSCE), Volume-3, Issue-4, September 2013
5. V.Ramachandran, et.al"Facial Expression Classification Systems with Emitional Back Propagation Neural Network " International Journal of Scientific and Engineering Research (IJSER) - (ISSN 2229-5518), Volume4, Issue 9.during Aug 2013
6. Lt. Dr. S. Santhosh Baboo, Lt. Dr. S. Santhosh Baboo," Face Emotion Analysis Using Gabor Features In Image Database for Crime Investigation", International Journal of Data Engineering (IJDE), Volume (2) : Issue (2) : 2011
7. P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 711–720, Jul. 1997.
8. S. Z. Li, X. W. Hou, H. J. Zhang, and Q. S. Cheng, "Learning Spatially localized, parts-based representation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2001, pp. 207–212.
9. L. L. Huang and A. Shimizu, "Combining classifiers for robust face detection," in Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2006, vol. 3972, pp. 116–121.

Authors Profile:

Vedantham Ramachandran, is a research scholar at Acharya Nagarjuna University, Nambur and an associate professor at Vasireddy Venkatadri Institute of Technology. He did B.Tech Computer Science & Systems engineering Degree from Andhra University and M.Tech in Computer Science Engineering from JNTU, Kakinada. He is very much interested in image processing, medical image retrieval, human vision & pattern recognition. He did several projects in image processing.

Dr. E. Srinivasa Reddy, Professor & Dean R&D in University Engineering College of Acharya Nagarjuna University. He did Master of Technology in Computer Science Engineering from sir Mokhsagundam Visweswariah University, Bangalore. He also did Mater of Science from Birla Institute of Technological Sciences, Pilani. He did Philosophical Doctorate in Computer Science Engineering from Acharya Nagarjuna University, Guntur. Prof.E.S.Reddy, guided several research scholars of PhD to success. His research areas of interest are in various domains including Digital Image Processing.

Mr. B. Sivaiahhas received the B.Tech. degree in 2007 in Computer Science & Engineering from Jawaharlal Technological University, Hyderabad, Andhra Pradesh, India. He received M.Tech. degreein 2010 in Computer Science & Engineering from Jawaharlal Technological University, Kakinada, Andhra Pradesh, India . Presently he is working as assistant professor in Department of Computer Science & Engineering, Vignan's University, Andhra Pradesh, India. His current research interests include image processing, neural networks, machine learning, intelligent systems, pattern recognition and artificial intelligence.