

Human Facial Emotion Recognition using Adaptive Sigmoidal Transfer Function in MLP Neural Network



Ismath Unnisa, Loganathan. R

Abstract: *The human face is very sensitive towards inner feelings particularly with different state of mind under various conditions. The facial expression has used in computer vision to understand the human response against stimuli. But the facial expression is also having the nature of variability and controllability hence its complete generalization from a computer vision point of view is very difficult and challenging, though acceptable performances can be achieved. In this paper, a two-stage based facial expression recognition model which carry the Principal component analysis as a feature extractor in the first stage and self-adaptive based activation function in feedforward neural network as a classifier in the second stage have applied. Use of principal component analysis reduces the dimension of features while the adaptive slope of transfer function provides another parameter along with weights to change in making learning faster and accurate. Six most dominant state of facial emotion like angry, surprise, sadness, normal, happy and fear have considered in this paper and performances have been tested over variable expressions. The benefit of the proposed model of self-adaptive activation function has verified through the benchmark XOR problem classification.*

Keywords: *Emotion, Facial emotion recognition, principal component analysis, neural network, MLP, Adaptive transfer function.*

I. INTRODUCTION

The role of facial expression or emotions in the human life is very important and natural because it shows what we feel exactly inside and without any verbal communication; this feeling can be observed through our facial emotions. Even appearances of emotions are certain extent can be controlled but complete control against the natural response is not possible. In social life, it plays a very important part to understand others against situations. Among the communication point of view, human communicate with each other either verbal approach or through different emotions state. It has been observed in different surveys that the contribution of nonverbal communication (dominated by facial expression) is two-third in compare to verbal communication which is one third.

Hence, there has been a lot of attention towards doing research in the area of facial emotion under the different stream of science and engineering like perceptual and cognitive sciences (behavioral science and in clinical practice), computer vision, computer animations, etc. From the engineering point of view, the most important factor for facial emotions is the design of automated recognition. For the human, recognition of any facial expression is generally effortless and quick, but machine point of view still it is a very challenging task. As there is variability and controllability exists with facial emotions, the required technology to support automated recognition must carry a high level of adaptiveness and intelligence. In result, the artificial intelligence approach particularly machine learning has become the one the most useful resource for developing the solution. Broadly in facial emotions recognition (FTR) there are three modules exist (i) Detection of the face and facial expression (ii) Feature extraction and transformation (iii) Classification of expression.

From an input image, image of face identified through various facial component like eyes, nose, mouth, eyebrow, etc. From these facial components, spatial and temporal features are extracted and if needed these features are transformed to the low dimension size. A classifier which has obtained the training over these features has applied to define the present emotion available at the input. In the past few years, there has been good progress observed in the area of face detection speed and feature extraction methods. Different classifier models having compact architecture and adaptiveness have been developed to extract the knowledge from these features (like neural network which has a reference of the human brain).

In this paper, an approach based on principal component analysis (PCA) and self-adaptive form of Multilayer perceptron feedforward Neural Networks for facial expression recognition. The input into our system is an image; then, proposed model predicts the facial expression label which should be one these labels among the six prominent part of the expression: anger, surprise, sadness, normal, happiness, fear.

II. RELATED WORK

There has been the number of research in the past over automated face recognition and having variances in their approach and methodology. A detailed survey of different possibilities of automated facial recognition has been presented in [1].

Revised Manuscript Received on October 30, 2019.

* Correspondence Author

Mrs. Ismath Unnisa¹, Department of Computer Science & Engineering, H.K.B.K College of Engineering, Bangalore, India, ismath.ayaan@gmail.com.

Dr. Loganathan. R², Department of Computer Science & Engineering, H.K.B.K College of Engineering, Bangalore, India, drloganathanr@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

The issues associated with a facial expression like facial motion and deformation extraction along with different classification methods have been explored in details. Possibilities to solve the facial expression intensity, expression dynamics and normalization have also covered. The application of deep learning like GoogLeNet and AlexNet for facial expression recognition has been applied in [2]. The eyes and mouth are the most dominant facial detection components available over the face.[3] has applied the local binary pattern to extract the features from these components to form the model for recognition of facial expression. Spontaneous emotions detection from facial expression has detected in [4].To detect human aggression, human face created using face detection mechanism and Gabor filter along with local binary pattern operator have applied to encode the face texture and finally support vector machine applied for classification purpose.[5] has proposed two-stage based facial expression recognition from front view face image which accommodate the facial features under large variation. The first stage provides the method for adaptively generating the initial model for Active Appearance Models (AAM) fitting and in second stage hybrid expression features introduced carrying geometric features, AAM shape, and appearance parameters. Based on system identification approach [6] using the Kohonen self-organizing map (KSOM) emotion recognition has proposed. KSOM has used geometric shape carried 26-dimensional features contained the information about eye, lip, and eyebrow. A cascaded structure of the neural network has applied in [7] for facial recognition considering the normalized image of the face as input and delivered face expression as the output. For each facial expression, an individual MLP has been applied. In the area of social psychology and social neuroscience particularly in regard of neuropeptide oxytocin [8] has presented a detailed study and analyzed how oxytocin influences the processing of emotion in faces by reviewing intranasal administration studies of automatic processing, selective attention, and emotion recognition. The convolutional neural network has applied in [9] to recognize the facial expression having integration of specific image pre-processing steps. Study over the use of Deep learning and convolution neural network has presented in [10]. Gradient-based encoding of facial component features has discussed in [11]. The purpose was to make recognition system sensitive with facial muscle deformation. Application OF facial emotion recognition has extended to command a mobile robot in [12]. Based on deep convolution neural network comprise of convolution layer and deep residual blocks six different emotions have been classified in [13].

III. PROPOSED METHOD

A. Adaptive Sigmoidal Transfer Function in MLP NN

In the learning process of the neural network through Backpropagation, the weights change rate of the neural network for the next iteration is proportional to the first-order derivative of the available nonlinear transfer function. Generally, the sigmoid function has applied with backpropagation learning because of soft limiting characteristics and having its derivative in simple function form of the sigmoid function itself. The sigmoid function can be defined as:

$$y = 1/(1 + e^{-px})$$

Where x is the input and p is the parameter responsible for the slope of the sigmoid function. The sigmoid function can be divided into the three regions as shown in Fig.1. The region R1 and R3 are representing the saturation region, where the change in function value does not change significantly from one point to other. Hence derivative at any point in this region has a small value. While region R2 having the characteristics of linearity and there is sharp change occurred from one point to other hence derivative is a large value. The input of the sigmoid function comes from the output of the linear combiner. If the output of linear combiner is either very high or very low the operation will shift in the saturation region of the transfer function in result the derivative will small which cause of very small change in the value of the weight. In result, a very small change in weights will cause of slow learning and a large number of iterations required.

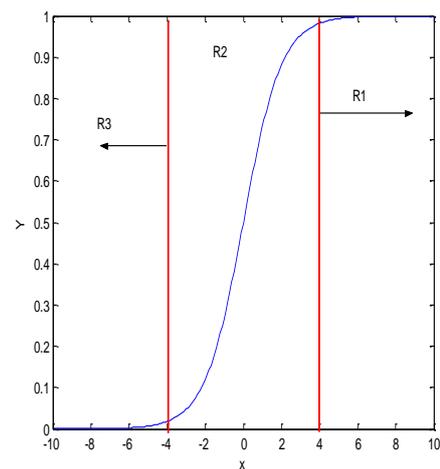


Fig.1. Sigmoid function operational region

The one possible remedy of this part is to decrease the slope of the transfer function as shown in Fig.2. to put the operation in the linear region. But this approach will lose the advantage of having multilayer network architecture. So there are two issues exist the first one, not having pre-information about the optimum slope of activation function and the second one all neurons should have the same slope or a different one. The difficulties associated with such a situation can resolve by providing the self-adaptiveness in the slope of the transfer function at each and every neuron independently so that optimal slope at every transfer function could appear for every iteration. This process will provide the optimal derivative value required at the point of the landscape to have the optimal change in the weight values. The optimum slope value can be achieved along with weight optimization simultaneously by minimizing the error function.

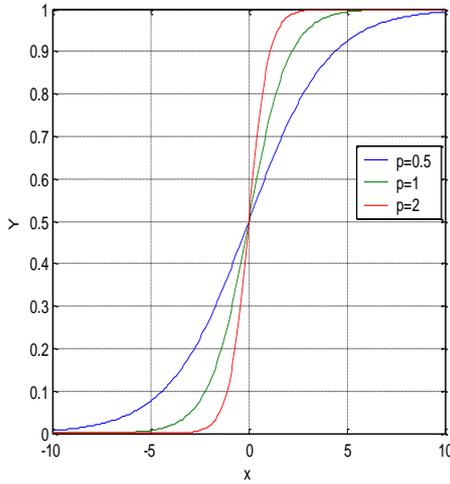


Fig.2. Effect over slope for different values of p

Specifically, the slopes are to be chosen so as to minimize the performance criterion

$$E_i = \frac{1}{2} (y_i - a_{out}^{(z)})^T (y_i - a_{out}^z)$$

$$= \frac{1}{2} \sum_{h=1}^n (y_{ih} - a_{out,h}^z)^2 \quad (1)$$

Where z denotes the number of layers in the network and $y_i \in \mathcal{R}^{n \times 1}$ and a_{out}^z are the derived and actual output, respectively of the network due to i^{th} training pattern. Consider an activation function of the sigmoid type given by Eq.2

$$f(u, p) = \frac{1}{(1 + e^{-pu})} \quad (2)$$

Where u is the input to the nonlinearity and p is the slope parameter which has to be adjusted so that Eq.1 has to minimize. Considering the nonlinearity of the m^{th} neuron in the z^{th} layer of the network, gradient approach can be applied by obtaining

$$p_m^z(t+1) = p_m^z(t) - \beta \frac{\partial E_q}{\partial p_m^z} \quad (3)$$

Using the chain rule, the second term on the right side in Eq.3 can be rewritten as

$$\frac{\partial E_i}{\partial p_m^z} = \frac{\partial E_i}{\partial u_m^z} \frac{\partial u_m^z}{\partial a_{out,m}^z} \frac{\partial a_{out,m}^z}{\partial p_m^z}$$

$$= -\delta_m^z \frac{1}{\frac{\partial a_{out,m}^z}{\partial u_m^z}} \frac{\partial a_{out,m}^z}{\partial p_m^z} = -\delta_m^z \frac{f_p(u,p)}{f_u(u,p)} \quad (4)$$

Where δ_m^z is the local error for the m^{th} neuron of the z^{th} layer, and $f_p(u, p)$ and $f_u(u, p)$ denote the partial

derivatives of the activation function with p and u respectively. Therefore, transfer function slope obtained as

$$p_m^z(t+1) = p_m^z(t) + \beta \delta_m^z \quad (5)$$

The momentum term has applied to make learning faster and stable.

B. Learning Algorithm with Adaptive Activation Function Slopes

1. Weight initialization has done through assigning a random number through uniform distribution in the range of [-1 to 1]
2. From the training data set, apply the training pattern and process the operation of architecture to obtain the network output.
3. Error at the present moment estimated by taking the difference between the observed outputs and expected actual output and define the local error as.

For output layer:

$$\delta_m^z = (y_m - a_{out,m}^z) g(u_m^z) \quad (6)$$

For hidden layer: $\delta_m^s = \sum_{h=1}^{n_2} \delta_h^{z+1} w_{hm}^{z+1} g(u_m^z)$ (7)

4. The change in the network weights defined as $w_{m,j}^z(t+1) = w_{m,j}^z(t) + \mu \delta_m^z a_{out,j}^z$ (8)
5. Upgradation of activation function slope defined as $p_m^z(t+1) = p_m^z(t) + \beta \delta_m^z + \alpha [p_m^z(t) - p_m^z(t-1)]$ (9)
6. Terminate the process if the no. of iterations equal to allowed values and network converged, otherwise go back to step 2.

IV. EXPERIMENTAL RESULT

A. XOR Classification

XOR classification problem is considered as a benchmark for the neural network hence the proposed algorithm has applied over it to evaluate the performance: The network of XOR problematic consists of two input nodes, two hidden nodes, and one output node. Performance of obtained result has shown in Table-I. It is clear that there is more accurate result have obtained by Adaptive Slope based architecture (ADSL) in compare to result obtained with Constant (Fixed) Slope of the activation function (FXSL), equal to 1. For the target '0' the outcomes from ADSL are around 2 times better in comparison to FXSL. For target '1', outcomes of ADSL is very close to 1 (0.9948) in compare to FXSL (0.9873). The total number of applied iterations was 5000 for both cases and from Fig3. it can observe that faster convergence achieved with ADSL. The final mean square error obtained in the learning process is also lesser. In the case of ADSL it is $1.0e-003 * 0.0307$ while for FXSL it is $1.0e-003 * 0.1796$.

Table-I : XOR Performance comparison

INPUT	TARGET	Achieved O/P (FXSL)	Achieved O/P (ADSL)	Error in O/P (FXSL)	Error in O/P (ADSL)
0 0	0	0.0119	0.0047	-0.0119	-0.0047
0 1	1	0.9873	0.9948	0.0127	0.0052
1 0	1	0.9873	0.9948	0.0127	0.0052
1 1	0	0.0159	0.0068	-0.0159	-0.0068
MSE		1.0e-003 *0.1796	1.0e-003 *0.0307		

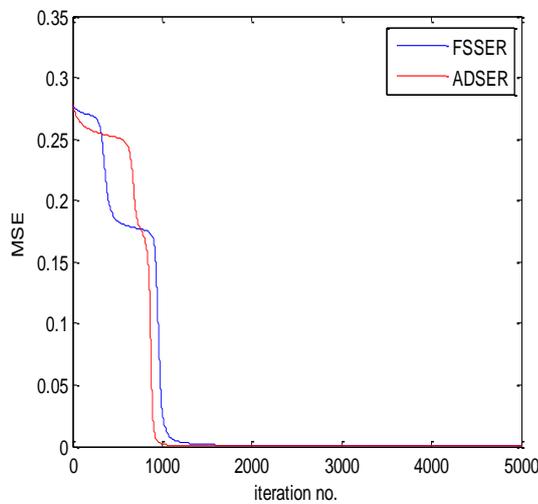


Fig.3 Convergence characteristic for XOR problem

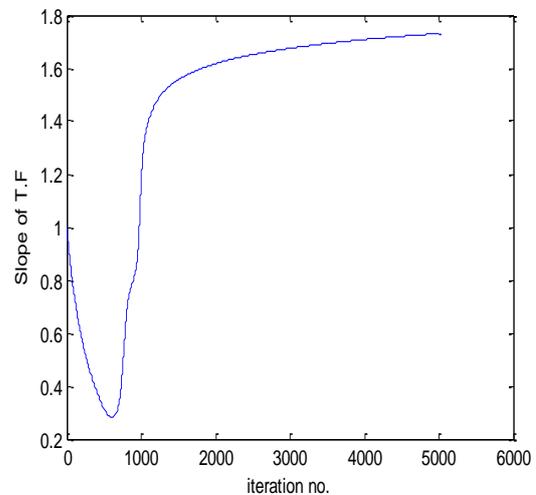


Fig.5 Slope variation in output node function

Table-II: Final slope value for XOR problem

Slope (hidden nodes) =	1.5803	1.5033
Slope(output layer nodes)=	1.7308	

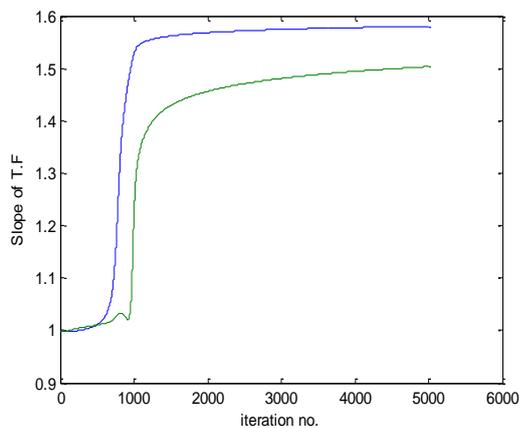


Fig.4 Slope variation in hidden nodes function

Adaptive characteristics in slope for hidden nodes and output nodes activation function have shown in Fig.4 and in Fig.5 and final iteration numeric values of slope have shown in Table-II. It is observed from the Fig.4 that, slope value for hidden nodes active function started to increase slowly from 1 at the beginning but after 1000th iteration there are sudden rise in the slope value. Characteristic for output node active function as appear in Fig.5, also pass through change of first decreasing than increasing. It is clear from both figures that there is very unusual kind of change has taken place in slope of active function for hidden as well as output nodes, which is not predictable. Hence rather than placing a fixed slope of with any value, it is always better to place the adaptive slope to make the learning faster.

Fig.6 shows the proposed model. For the various different state of face, first 3 principal components have considered and linearization has applied to transform the all the values of 3 PCA components into one-dimensional array so that all information available in 3 PCA components can be applied simultaneously to the neural network. The normalization has applied to keep the values of each element of the array in the range of [-1 1]. This will make the initial operation in the active region of the sigmoid function. Three-layer of MLP architecture has applied where the input number of nodes is (256*3) and output nodes were 6, each one belongs to represent each class of emotion. For a particular facial emotion class, the particular output nodes output will be maximum one while remaining output nodes will be minimum zero. The normalized information for each image has learned using the gradient descent algorithm. In each iteration along with hidden and output layer weights, the slope of all active nodes in hidden and output layer have modified. The target for each emotion class has shown in Table-III.

Table-III : Target values corresponding to different Facial Emotion

Emotion	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6
Anger	1	0	0	0	0	0
Surprise	0	1	0	0	0	0
Sadness	0	0	1	0	0	0
Normal	0	0	0	1	0	0
Happiness	0	0	0	0	1	0
Fear	0	0	0	0	0	1

In the experiment, the different values of hidden layer nodes 20, 40, 50 and 60 have tested to get the comparative learning performances. The learning rate and momentum constant have kept as 0.2. The obtained performances with different hidden nodes have shown in Fig.7. It is clear the learning was faster with 40 hidden nodes hence opted.

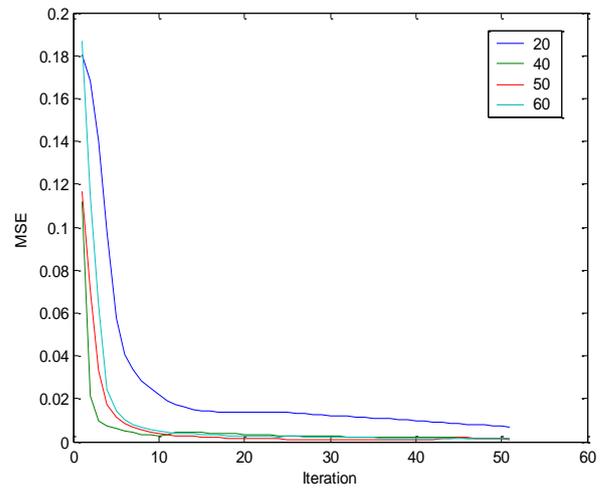


Fig7. The learning Convergence with different Hidden nodes.

B. Facial emotion Recognition

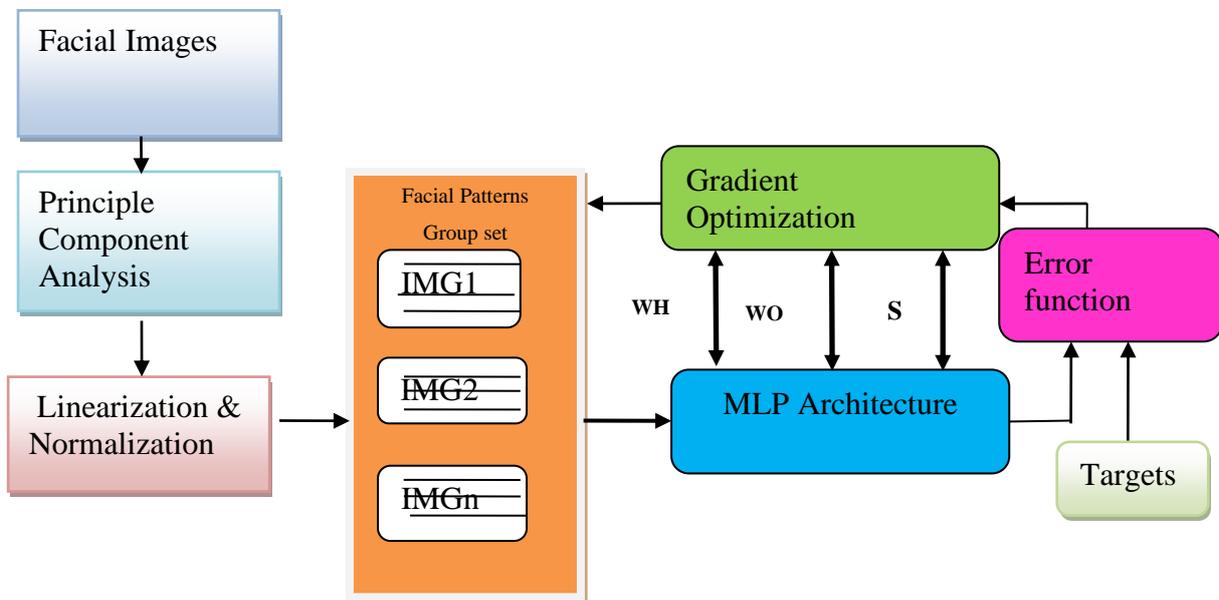


Fig.6. Functional block diagram for Facial Emotion Recognition

With the various different facial emotion images as shown in Fig.8 Training has applied. The allowed number of iteration is when MSE is greater than 0.001. The performance over the training images has shown in Fig.9. It is clear that there was sharp decision confidence in all the cases. The numeric values for each class have shown in Table-IV. The obtained test images decision confidence has shown in Fig.10. It can observe there were significant changes in the facial emotion even after there is good accuracy has achieved with very less training data set.

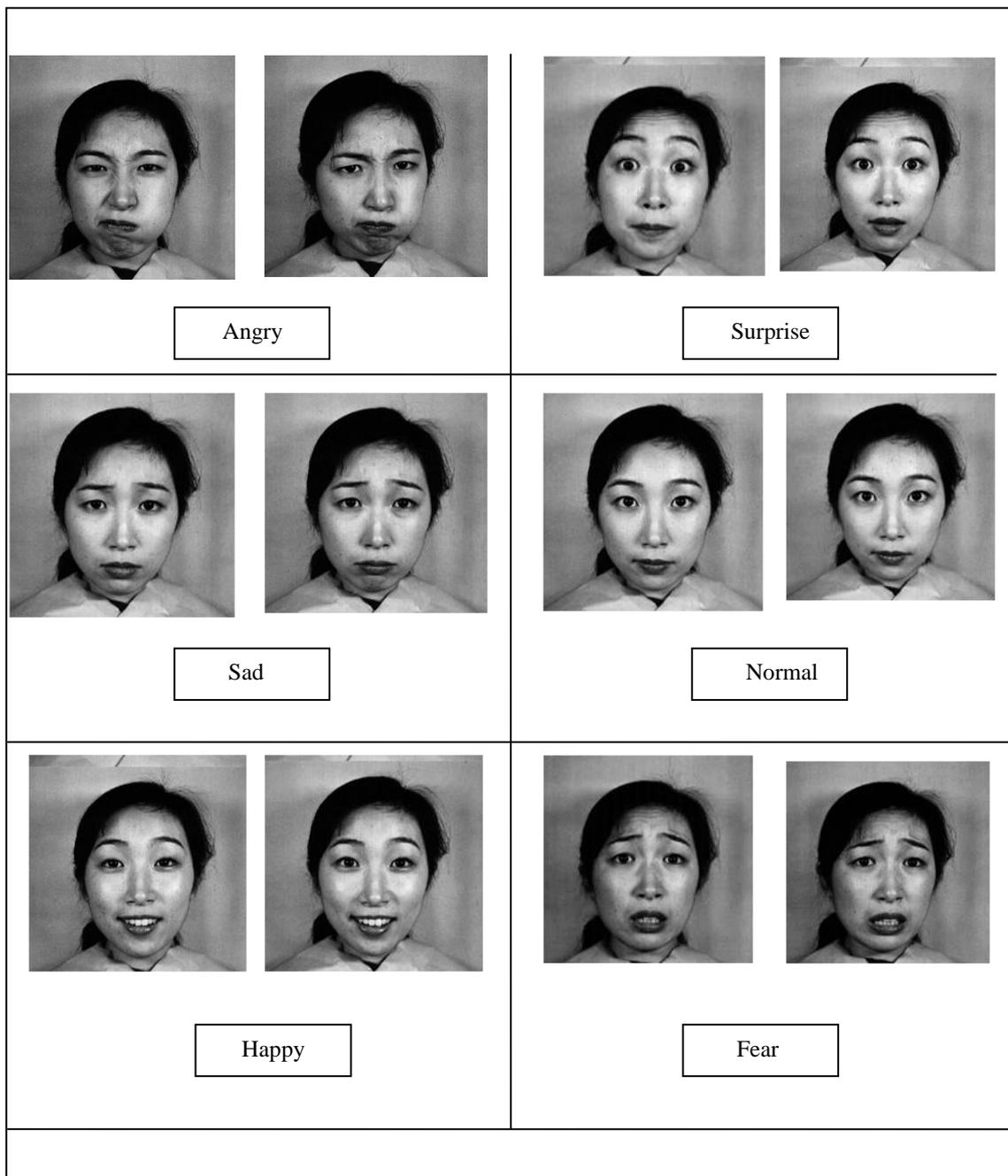


Fig.8. Different Facial Emotion Images for Training

Table-IV: Decision Confidence in Training and Test cases

Input Emotion	Angry	Surprise	Sad	Normal	Happy	Fear	Recognised Emotion
Angry							
Tr.	0.9859	0.0095	0.0025	0.0000	0.0088	0.0144	Angry
	0.9796	0.0114	0.0170	0.0091	0.0006	0.0007	Angry
Test	0.9916	0.0215	0.0010	0.0000	0.0052	0.0379	Angry
Surprise							
Tr.	0.0012	0.9860	0.0088	0.0001	0.0002	0.0005	Surprise
	0.0101	0.9846	0.0000	0.0085	0.0114	0.0000	Surprise
Test	0.0548	0.4647	0.0001	0.0008	0.4202	0.0000	Surprise
Sad							
Tr.	0.0065	0.0036	0.9822	0.0107	0.0077	0.0068	Sad
	0.0153	0.0019	0.9807	0.0118	0.0085	0.0105	Sad
Test	0.0039	0.0222	0.9116	0.1435	0.0000	0.8362	Sad
Normal							
Tr.	0.0002	0.0073	0.0167	0.9880	0.0015	0.0003	Normal
	0.0002	0.0092	0.0063	0.9852	0.0007	0.0005	Normal
Test	0.1058	0.2513	0.0002	0.0000	0.0259	0.0008	Surprise
Happy							
Tr.	0.0059	0.0019	0.0062	0.0043	0.9878	0.0005	Happy
	0.0006	0.0155	0.0003	0.0004	0.9857	0.0002	Happy
Test	0.0108	0.1927	0.0004	0.0003	0.4120	0.0008	Happy
Fear							
Tr.	0.0126	0.0064	0.0009	0.0051	0.0000	0.9857	Fear
	0.0087	0.0001	0.0061	0.0059	0.0069	0.9886	Fear
Test	0.0639	0.0114	0.0039	0.0024	0.0000	0.8996	Fear

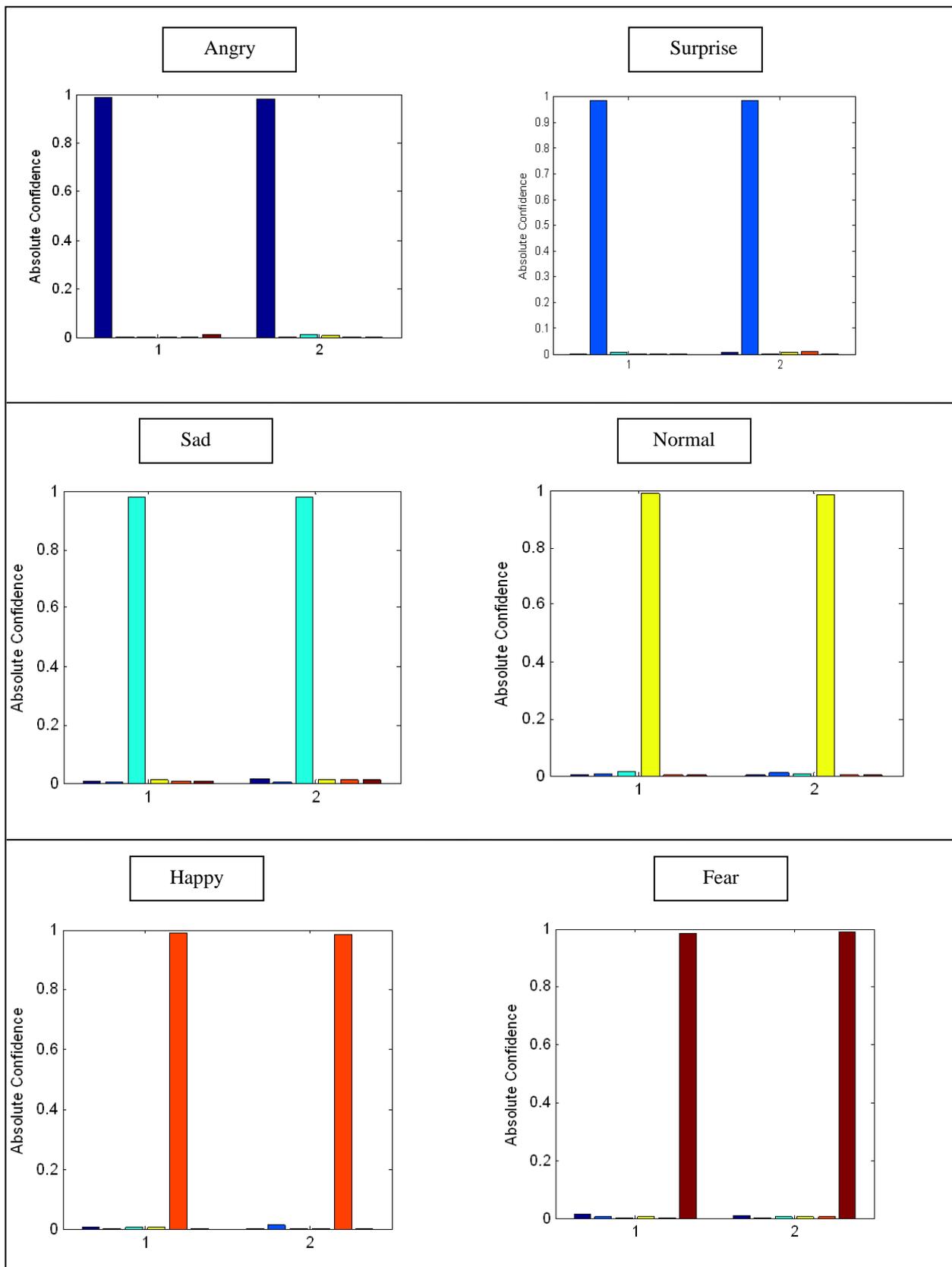
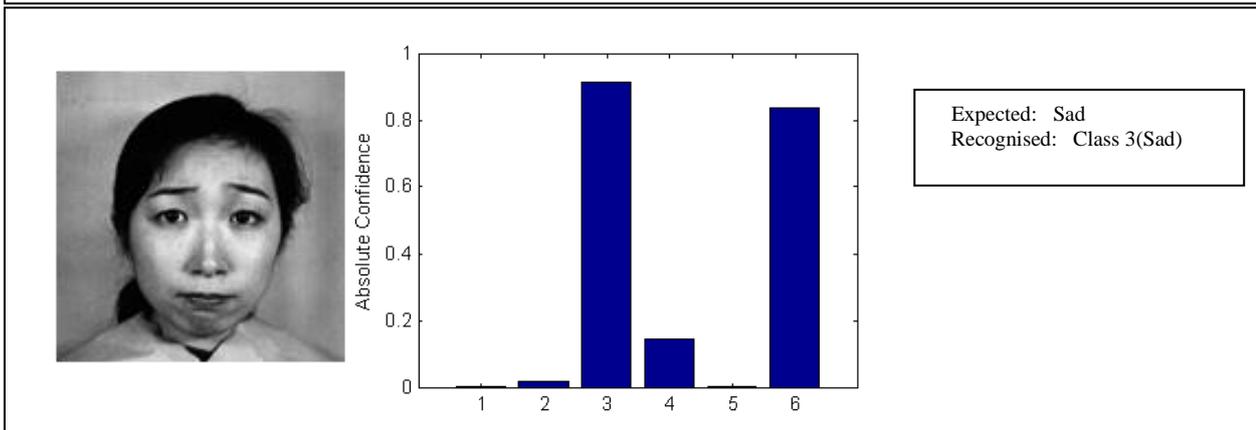
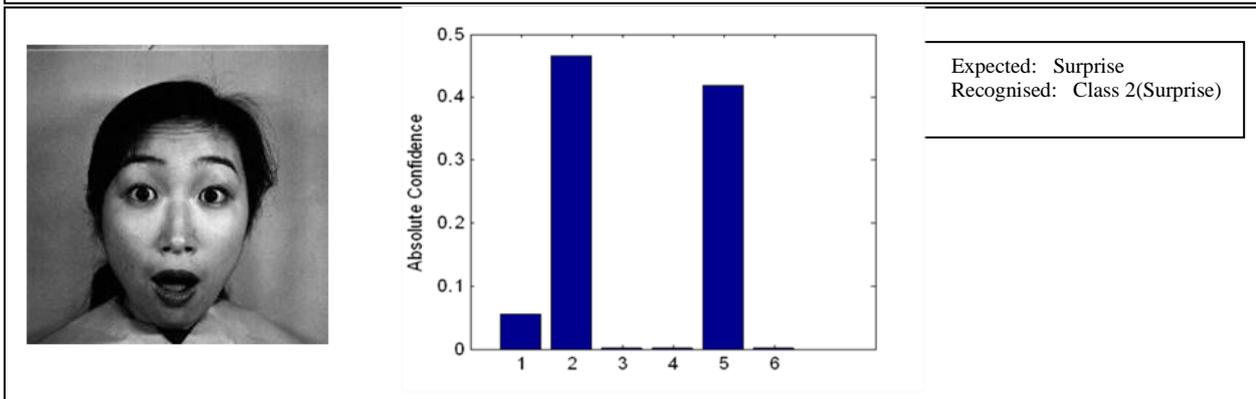
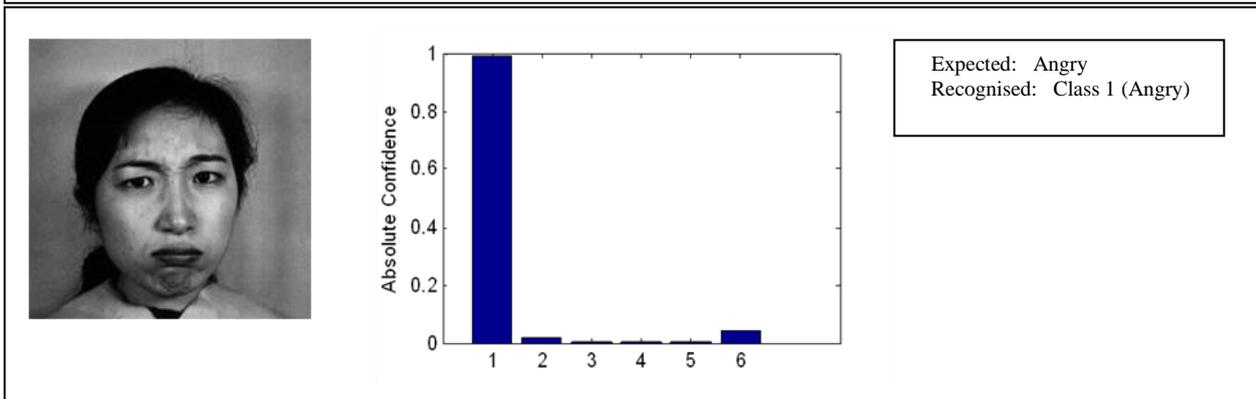
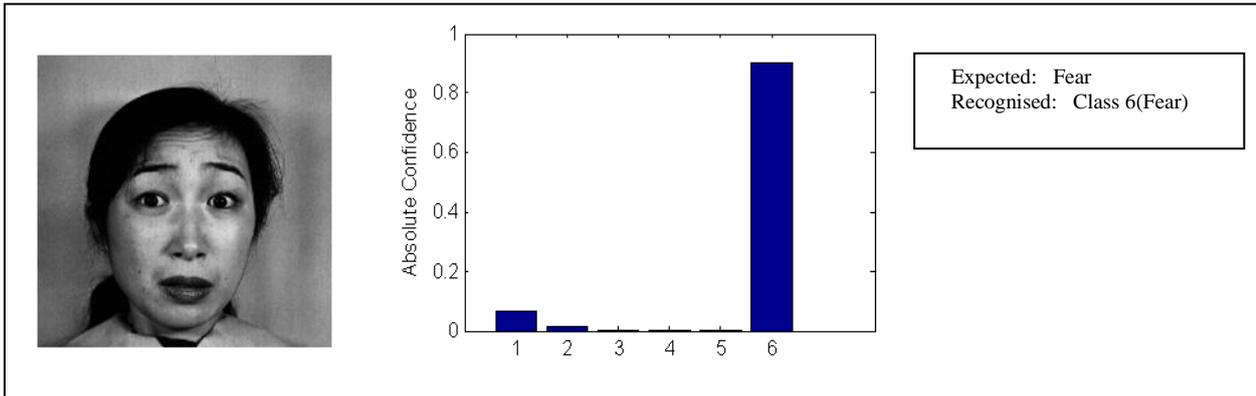


Fig.9. Decision Confidence in Training Images.



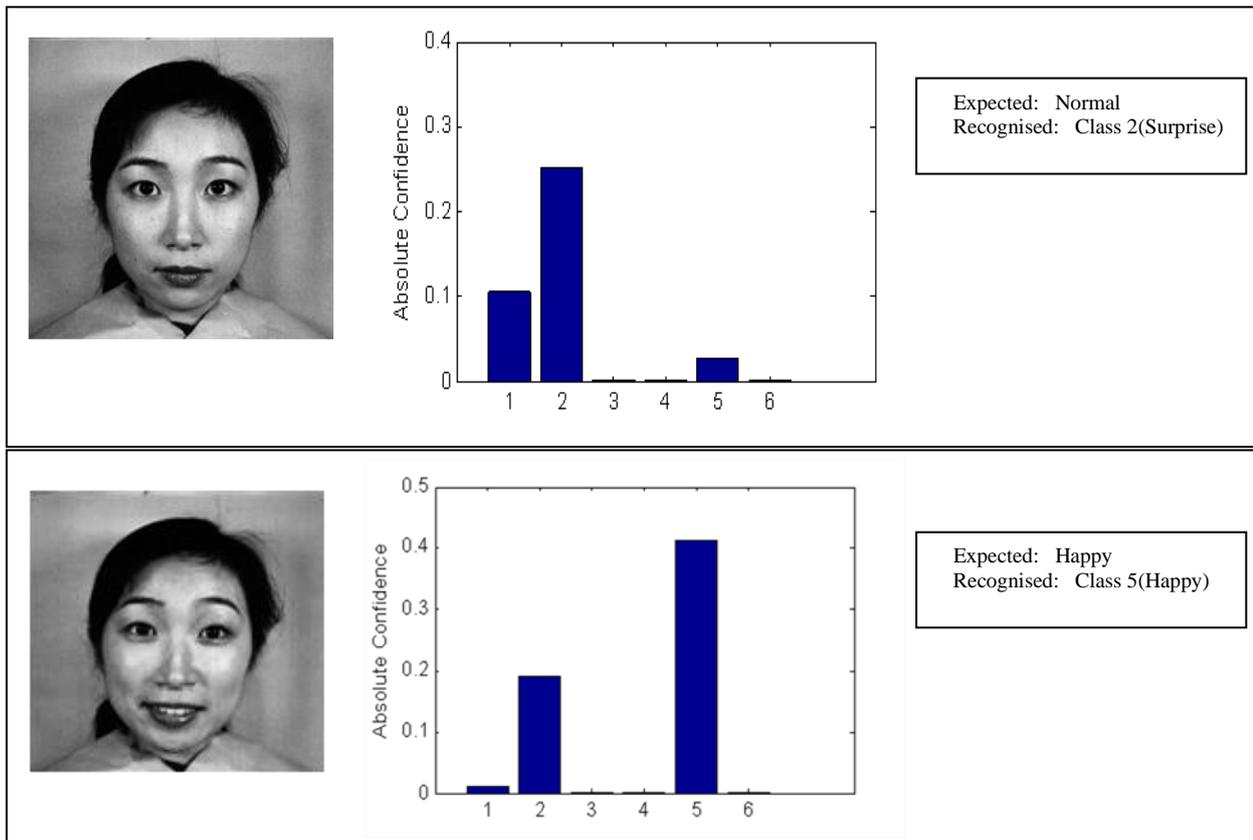


Fig.10. Different Test Facial Emotion and their Recognition.

V. CONCLUSION

The challenge of handling facial emotion recognition in an automated manner has proposed in this work. The proposed method has simplicity in design and implementation. The feature extraction has optimized through the use of PCA which helped to extract the feature variation in the emotion available in the faces. The classifier quality has enhanced by making the adaptiveness in the transfer function which helped to learn the things faster. In this way now along with weights one more dimension available to adjust the parameters to meet the target objectives. There are possibilities to improve the performance further by the use of extensive data set to capture various other variation in emotions appeared in the same category.

REFERENCES

1. B.Fasel ,JuergenLuettin, "Automatic facial expression analysis: a survey", Pattern Recognition, Volume 36, Issue 1, January 2003, Pages 259-275
2. Giannopoulos P., Perikos I., Hatzilygeroudis I. (2018) Deep Learning Approaches for Facial Emotion Recognition: A Case Study on FER-2013. In: Hatzilygeroudis I., Palade V. (eds) Advances in Hybridization of Intelligent Methods. Smart Innovation, Systems and Technologies, vol 85. Springer, Cham
3. Wang W., Chang F. (2014) Facial Expression Recognition Based on Feature Block and Local Binary Pattern. In: Pan L., Păun G., Pérez-Jiménez M.J., Song T. (eds) Bio-Inspired Computing - Theories and Applications. Communications in Computer and Information Science, vol 472. Springer, Berlin, Heidelberg
4. Ewa Piątkowska, Jerzy Martyna, "Facial Expression Recognition for Detecting Human Aggression", International Conference on Artificial Intelligence and Soft Computing ,ICAISC 2012: Artificial Intelligence and Soft Computing pp 574-582.

5. LingZhang, SipingChenTianfuWangZhuoLiu, "Automatic Facial Expression Recognition Based on Hybrid Features", Energy Procedia, Volume 17, Part B, 2012, Pages 1817-1823.
6. AnimaMajumderLaxmidharBeheraVenkatesh K.Subramanian, "Emotion recognition from geometric facial features using self-organizing map", Pattern Recognition, Volume 47, Issue 3, March 2014, Pages 1282-1293.
7. Kukla E., Nowak P. (2015) Facial Emotion Recognition Based on Cascade of Neural Networks. In: Zgrzywa A., Choroś K., Siemiński A. (eds) New Research in Multimedia and Internet Systems. Advances in Intelligent Systems and Computing, vol 314. Springer, Cham
8. Ellenbogen M.A. (2017) Oxytocin and Facial Emotion Recognition. In: Hurlmann R., Grinevich V. (eds) Behavioral Pharmacology of Neuropeptides: Oxytocin. Current Topics in Behavioral Neurosciences, vol 35. Springer, Cham
9. André TeixeiraLopes, Alberto F.DeSouza, "Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order", Pattern Recognition, Volume 61, January 2017, Pages 610-628
10. Byoung Chul Ko, "A Brief Review of Facial Emotion Recognition Based on Visual Information", Sensors (Basel). 2018 Feb; 18(2): 401. .Published online 2018 Jan 30. doi: 10.3390/s18020401 .PMCID: PMC5856145m, PMID: 29385749
11. Malika Arora, Munish Kumar, "Facial Emotion Recognition System Based on PCA and Gradient Features", National Academy Science Letters, December 2018, Volume 41, Issue 6, pp 365-368 .
12. Chérifa Z., Nasr Eddine B. (2018) Facial Expressions Recognition: Development and Application to HMI. In: Amine A., Mouhoub M., Ait Mohamed O., Djebbar B. (eds) Computational Intelligence and Its Applications. CIAA 2018. IFIP Advances in Information and Communication Technology, vol 522. Springer, Cham
13. DeepakKumarJain, PouryaShamsolmoali, ParamjitSehdev, "Extended deep neural network for facial emotion recognition", Pattern Recognition Letters, Volume 120, 1 April 2019, Pages 69-74.

AUTHORS PROFILE



Mrs. Ismath Unnisa is a Research Scholar in Computer Science & Engineering at HKBK College of Engineering, VTU, Belagavi with Research title “Hybridization of Face and Associated Emotion Recognition using Evolved Neural Network”. I have 8 years of Teaching Experience and 6 Months of Research Experience. Presently, Working as an

Assistant Professor in Computer Science & Engineering at Islamiah IT, Bangalore. I completed my M.Tech from The Oxford College of Engineering, Bangalore, in the year 2012. I have presented and published 11 papers at National and International Journal and also awarded second prize for the paper presented at M.S.Ramaiah Institute of Technology, Bangalore. My current interests include image processing and analysis, Facial recognition, Emotional Recognition, computer vision and pattern recognition.



Dr. Loganathan R received the Ph.D degree in Computer Science & Engineering from the Sathyabama University, Tamilnadu for the thesis entitled “A Novel Medical Image Compression Technique with Lossless Region of Interest using Adaptive Active Contour” in the year 2014. I have 20 Journal published Papers and 8 conferences presented and Published for National and International Journal. I am guiding 6 members for pursuing

their Ph.D. I am having 24 Years of Experience in Teaching. I am working at HKBK College of Engineering since 1995. Presently, designated as Professor & Head of Computer Science & Engineering Department. I am also member under professional bodies for ISTE and CSI. My current interests include image processing and analysis, Facial recognition, Emotional Recognition, computer vision and pattern recognition.