Facial Expression Detection Based On Local Binary Pattern and Back Propagation Neural Network

Mousa Kadhim Wali, Md. Hussein Baqir, Majid S. Naghmash

Abstract* — Facial expression has always been used to show human's feeling. There are many types of human's emotion. Nevertheless, there are six primary emotions are shown in the similar way by people throughout the world regardless of culture, which are sadness, anger, happiness, fear, disgust and surprise. Since facial expressions are universal, therefore utilizing them to create several applications is possible. Face detection is crucial process where it would directly affect the emotion detection accuracy. So, this work utilizes Open CV implementation of Viola-Jones face detection library to detect faces automatically using Japanese Female Facial Expression database. Image processing and emotional classification was done in Matlab since it has excellent support tools for image processing and neural network training. Four Features namely; contrast, correlation, energy, and homogeneity were extracted based on Gray-Level Co-occurrence Matrix method after preprocessing by histogram and adaptive filter. Back propagation neural network been used in this research which yield of 87.5 % detection accuracy. A Graphical User Interface (GUI) was developed using Graphical User Interface Development Environment in Matlab.

Keywords: Facial, BPNN, LBP, GLCM

I. INTRODUCTION

This research has significance in many fields such as medical, device developer, robotics and Human-Computer Interface researches. As an example, robot can display the 'happy' feeling when it detects a face (through camera) or it hears a voice. As for the medical field, this system can be used to constantly analyze patients face for any kinds of fatigue or pain signs and alert the doctors remotely since it utilizes wireless devices. This system can help to develop application for physically disabled persons such as a door that can be open when it detects a happy face and so on. Other applications would be intelligent entertaining systems for kids, interactive computers, intelligent sensors and social robots.

Facial expression has always been used to show human's feeling. There are unlimited types of human's emotion.

Nevertheless, there are six primary emotions are shown in the similar way by people throughout the world regardless of culture and they are namely; sadness, anger, happiness, fear, disgust and surprise [1].

The face can be represented in different ways either as a whole unit based on *holistic representation or a*s a set of features based on a*nalytic representation* [2].

Human face detection and recognition is one of the most active research fields which interfere with other fields such as image processing, pattern recognition, and computer vision. Among the emotion detection methods, facial expression reaches up to more than 85% accuracy and considered as the best detector.

It's followed by physiological method which achieved 80% accuracy based on speech and the bimodal (speech and facial), got 72% percent [3][4].

For feature extraction and detection, there are many methods. Face detection based on Viola-Jones algorithm which uses features similar to Haar wavelets in gray scale intensities [5].

Recently, Mahesh et. al (2011) utilizes Principal Component Analysis (PCA) for facial expression detection and manage to get 86% percent accuracy [6]. Ganesh et al., (2003) used Gabor wavelets for feature extraction [7]. While, Local Fisher Discriminant Analysis feature extractor was used by [8], recently, Local Binary Patterns (LBP), which originally proposed for texture analysis has been successfully applied as a local feature extraction method in facial expression recognition [9] [10].

II. METHODOLGY

This work uses Japanese Female Facial Expression (JAFFE) [11] database which is freely available in internet. This database contains more than 200 images of Japanese females who portraits 7 facial expressions namely angry, happy, sad, surprise, neutral, fears and disgust.

This research utilized four facial expressions which are angry, happy, sad, and surprise. So, total image used in this project is reduced to 120 due to this reason. The whole methodology is explained in details in the flowchart of fig. (1).

Manuscript received March, 2014.

Mohammed Hussein Baqir. was born in 1956 in Iraq - Baghdad holds a Baklores engineering Electronic and Communications in 1976 and master's degrees in engineering electronic Technology University in Baghdad.

Mousa Kadhim Wali received the B.Sc. degree in Electrical Engineering from Baghdad University, Iraq in 1985, the M.Sc. degree in Electronic Engineering from Bradford University, U.K.

Majid S. Naghmash Received His Msc. degree in the satellite communication engineering from University of Technology Baghdad – Iraq

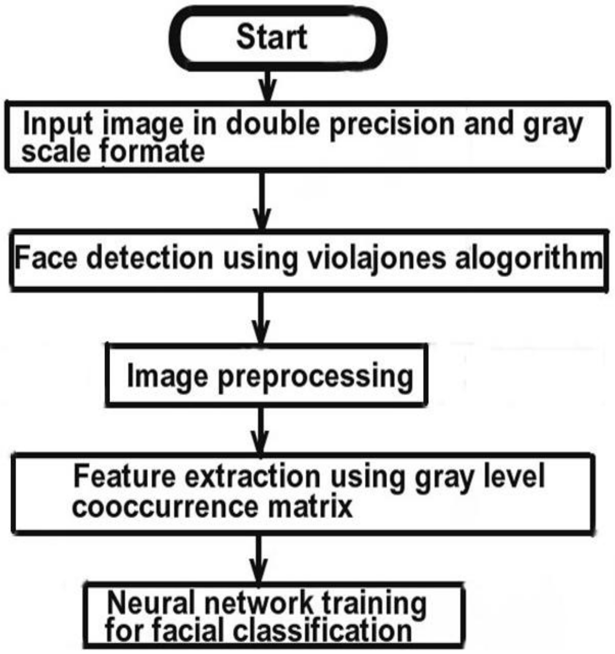


Fig. 1 The whole methodology flowchart

III. FACE DETECTION USING VIOLA-JONES ALGORITHM

Once images are obtained, Face detection and separation of facial images from the background has been implemented using viola-Jones algorithm which is one of the most robust face detection methods.

It is essentially a rapid object detection algorithm, uses monochromatic information to compute the features. Basic principal of this algorithm is to scan a sub-window (detector) which is capable of detecting faces across a given input image. Implementation of Viola-Jones algorithm which is fast and efficient is given in figure 2 as a flowchart of Viola-Jones method [12].

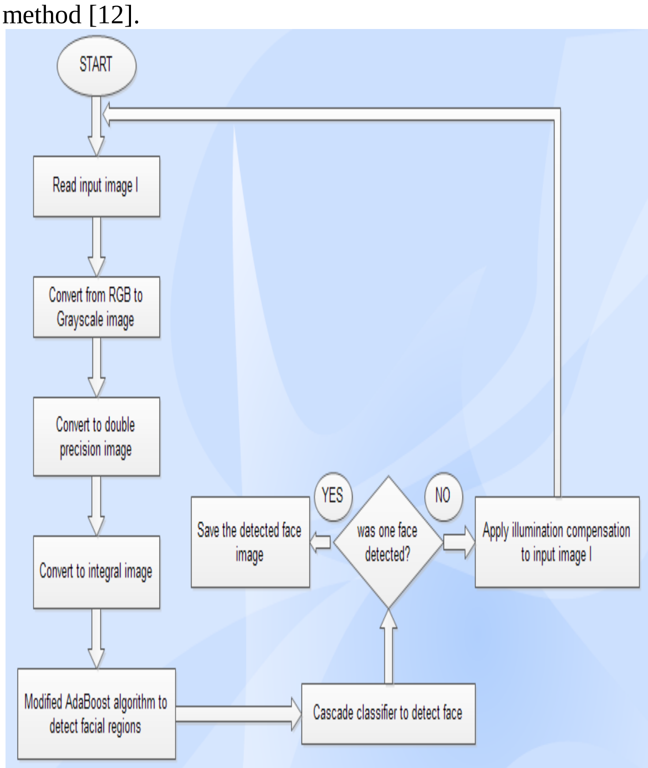


Fig. 2 Flow chart for Viola-Jones algorithm

First step is to convert the input image into an integral image as shown in figure 3. Then, each pixel is made equal to the entire sum of all pixels above and to the left of cornered pixel. This enables calculation of any given rectangular area within the desired region using only four values. Figure 3 shows calculation of integral image.

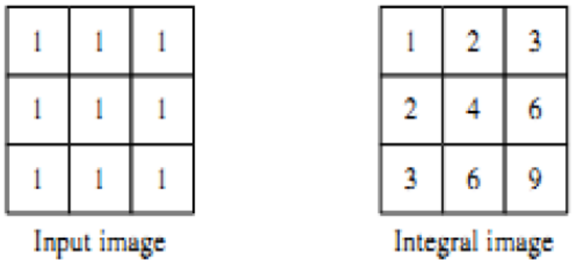


Fig. 3 Integral image computation

Once features are extracted, modified "AdaBoost" algorithm is used by Viola-Jones to detect features which are corresponds to face region. This is due to the fact that, feature value is high if it is extracted from facial region than non facial regions. "AdaBoost" algorithm constructs strong classifier based on weighted combination of weak classifiers. This algorithm is modified to select only the best features since only a small number of features are expected as an input to the classifiers.

The first feature takes advantage of the difference in intensity between the region of the eye and the region of the upper cheek. The second feature calculates the difference between the two eyes and the center bridge of the nose as shown in figure 4.

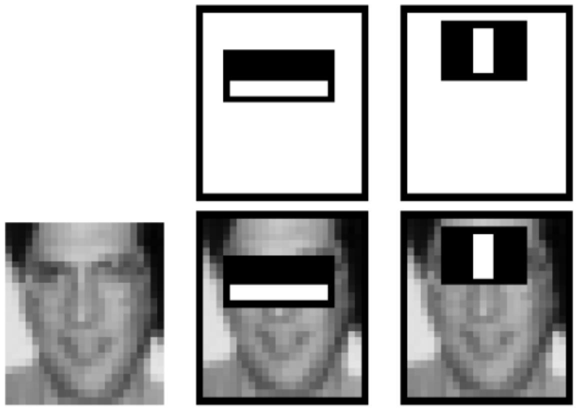


Fig. 4 First and second features obtained by modified AdaBoost algorithm.

Actually, cascade classifier composed of multiple stages, where each stage has strong classifier.

A positive result from the first classifier triggers the evaluation of a second classifier which has also been adjusted to achieve very high detection rates.

A positive result from the second classifier triggers a third classifier, and so on as shown in fig. 5.

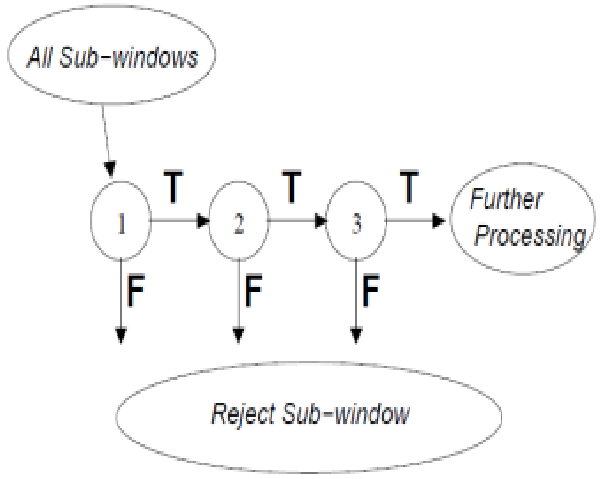


Fig. 5 The cascade classifier [5]

Implementation of this algorithm was done in Matlab with the help of Open CV (Open Source Computer Vision) library. This Open CV library already contains a trained classifier which can be used directly to detect frontal faces.

IV. PREPROCESSING

Once face detection has been done on the database, the following step is to do feature extraction from the image. Before this step, image preprocessing should be done to enhance the image.

First, we have to remove noises from image as much as possible since it would directly affect the neural network accuracy. Therefore, each image was filtered using Adaptive Filtering method provided in Matlab Image Processing Toolbox. This method claimed to preserve the edges in images (The Mathwork, 2005) based on "*wener2*" function provided by Matlab to filter the images adaptively by acclimating local image variance. Level of smoothing needed by the image is inversely proportional to image variance. Then, Contrast-Limited Adaptive Histogram is utilized to adjust and enhance the contrast in our filtered face image.

Histogram that defined above contains information about distribution of edges, flat area and spots for a whole image. But to create an excellent face representation, spatial information also must be collected together as feature. Hence the face image is divided into *m* small regions R_0, R_1,……..,R_m and a spatially enhanced histogram is defined as:

$$H_{i,j} = \sum_{x,y} I(f_l(x,y)) \text{ -------------------------------- (1)}$$

Where $I(x,y) \in R$

$i = 0, \ldots n-1, j = 0, \ldots, m-1$**--------------------------(2)**

n is the number of different labels produced by the LBP operator and.

$$I(A) = \{ {}^{1\ A\ is\ true}_{0\ A\ is\ false} \text{ ------------- (3)}$$

This last step converts our filtered image into equalized image. There are many parameters which can be changed to get an optimal image such as 'clip limit' which defines the contrast level and 'distribution' which defines the contrast level distribution on image. In this research two parameters were adjusted as follows: for 'clip limit', 0.01threshold is used and for 'distribution', the 'uniform' function is applied.

V. FEATURE EXTRACTION USING GLCM

In this research the features extraction based on Gray-Level Co-occurrence Matrix (GLCM) features. Using built-in "*graycomatrix*" function in Matlab Image Processing Toolbox, we extracts contrast, energy, correlation and homogeneity features of an image.

Is it said that GLCM can reveal certain properties about the spatial distribution of the gray levels in the texture image (Matlab 2011).

GLCM is created as illustrated below in figure 6.

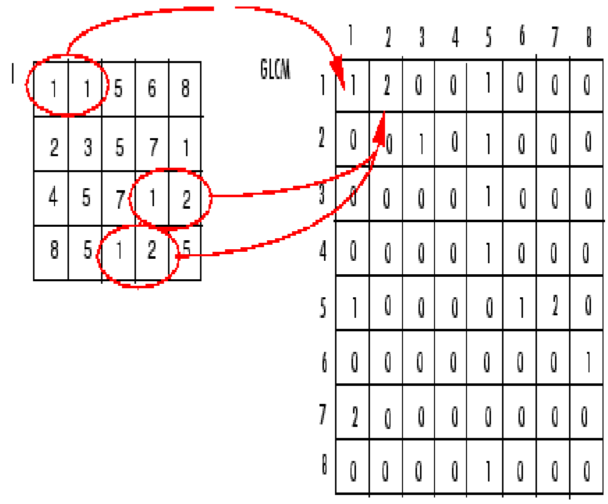


Fig. 6 GLCM matrix

Figure 6 shows how "*graycomatrix*" calculates the values in a GLCM. In the output GLCM, element (1,1) contains the value 1 because there is only one instance in the input image where two horizontally adjacent pixels have the values 1 and 1, respectively. GLCM (1,2) contains the value 2 because there are two instances where two horizontally adjacent pixels have the values 1 and 2 and so on. "*graycomatrix*" continues processing the input image, scanning the image for other pixel pairs (*i,j*) and recording the sums in the corresponding elements of the GLCM. After GLCM is created, we can derive contrast, energy, homogeneity and correlation using "*graycoprops*" function. Table 1 below describes the meaning of above mentioned features.

Table 1 Description of features in GLCM

Feature	Description
Contrast	Measures the local variation in GLCM
Correlation	Measures the joint probability occurrence of the specified pixel pairs
Energy	Sum of squared elements in GLCM
Homogeneity	Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal

GLCM features were extracted from 120 images in our study. It comprises of 30 angry images, 31 happy images, 30 sad images and 29 surprise images. All the extracted features were stored in respective Microsoft Excel file for further processing.

VI. NEURAL NETWORK

Neural networks are trained to perform a particular function by adjusting the values of the connections (weights) between elements (Klimis Symeonidis, 2000) [13].

Usually neural networks are trained, so that a particular input matches a specific target output.

Target of the input must be set priori to training. Neural networks have been applied successfully in solving complex problems like in field of bioinformatics, forecasting, pattern recognition, communications, and robotics and control systems [14].

The Multi-Layer Perception (MLP) is the most popular network architecture used today since many problems are non-linear.

Each unit performs a biased weighted sum of their inputs and passes this activation level through transfer function to produce their output. The units are arranged in a layered feed forward topology.

The network thus has a simple interpretation as a form of input-output model, with the weights and thresholds (biases). Such networks can model functions of almost arbitrary complexity, with the number of layers, and the number of units in each layer, determining the functions complexity.

Important issues in MLP design include specification of the number of hidden layers and the number of units in these layers (Bishop, 1995) [15].

The most common learning algorithm for MLP is 'back-propagation' or 'backprop' algorithm.

According to Haykin (1999) [16], the essence of back-propagation learning is the encoding of input-output mapping into the synaptic weights and the thresholds of MLP as shown in fig. (7).

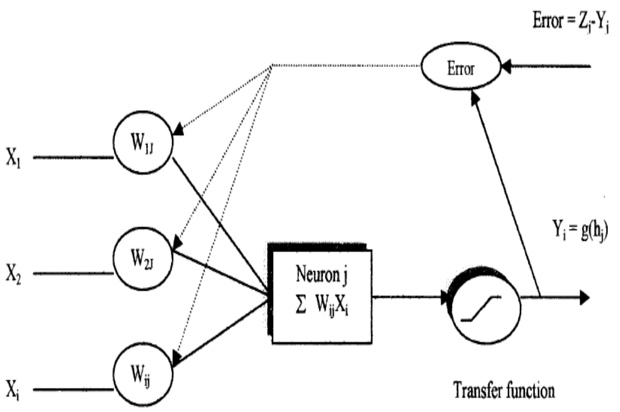


Fig. 7 Back propagation network

Training and testing of a neural network involves multiple stages as shown figure 8. It starts with determining best hidden unit, and then followed by selecting best learning rate, momentum rate, activation function and finally maximum number of epoch.

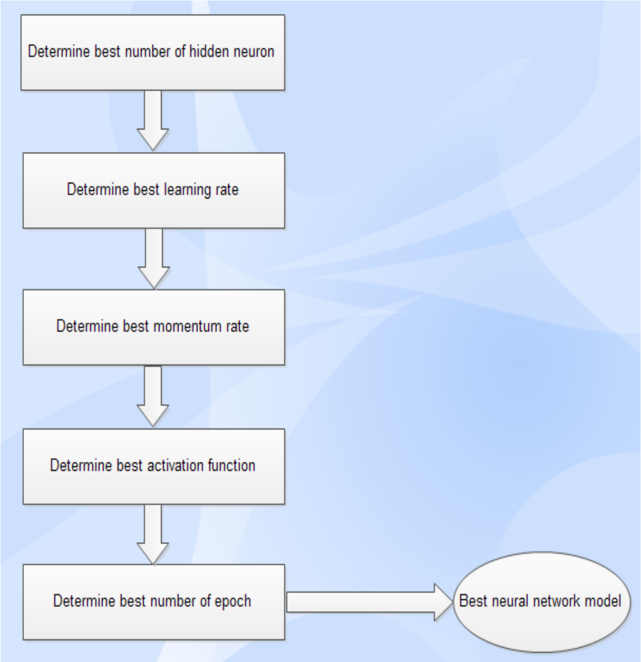


Fig. 8 Stages to get best neural network model

Initial parameters need to be determined in order to obtain best neural network through series of iterations or process. Among the initial parameters, number of epoch plays an important role because if the epoch number is too high the network would over fit and gives poor generalization performance. On the other hand, if the number of epoch is too low this let the network under fit and wouldn't learn. So, several experiments were carried out in order to determine the number of initial epoch.

The equation used to evaluate and measure the performance of NN is the prediction accuracy in conjunction with a confusion matrix that is given by keally (1999) [17].

$$Accuracy,\% = \frac{\sum_{i=1}^{N}[Number\ of\ Correct\ detection\ of\ Class\ i]}{Total\ Number\ of\ Samples} x100\%$$

-------------------------- (4)

Where N is the number of classes

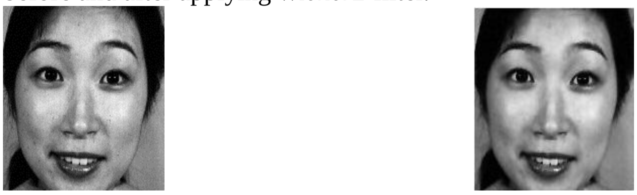
VII. RESULTS

Viola-Jones implementation in Matlab for face detection yields very good results. Figure 9 shows sample image from JAFFE database and detected face using Viola-Jones classifier with 100 % accuracy of detection face image.

Fig. 9 Sample image from JAFFE database and detected face

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

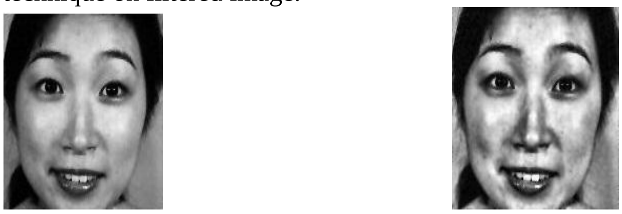
Before applying GLCM on the database images, images had been preprocessed so a better descriptor can be formed later. Figure 10 below shows sample image from JAFFE database before and after applying *Wiener2* filter.



Before **After**

Figure 10 Sample image before and after *wiener2* filter application.

As we can see in the image, wiener2 filter reduces noise on image and smoothen it. This is vital since image with noise would affect NN generalization capability. Once filter has been applied, next step would be applying adaptive-histogram technique to increase the contrast of the gray scale image. Figure 11 shows result of applying adaptive histogram technique on filtered image.



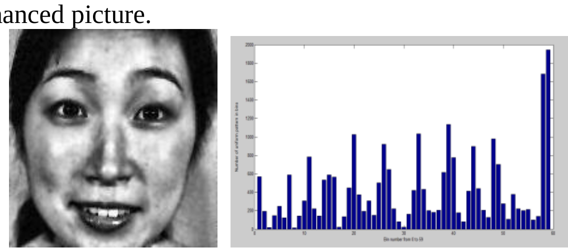
Before **After**

Figure 11 Applying adaptive histogram techniques

As we can see, adaptive histogram technique has increased the contrast of the image with noticeable changes. These changes are vital since even small change can make difference on NN accuracy later.

Next phase is feature extraction phase. During this phase, we had extracted GLCM's 59 feature vector which represent the number of uniform and non-uniform binary pattern in test image.

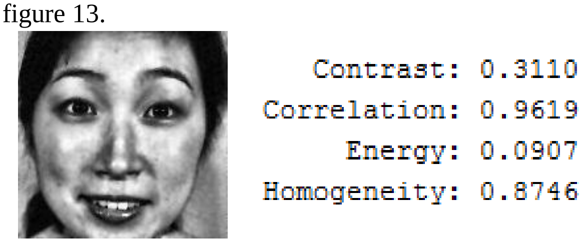
Figure 12 shows GLCM features in bar form for the contrast enhanced picture.



Input image **Uniform pattern in bins**

Figure 12 Input image and number of patterns in bins

This GLCM outputs four features for each image as given in figure 13.



Input image **GLCM features**

Fig. 13 Input image and snapshot of GLCM features from Matlab

As the final result of this study, for best neural network model is given in table 2. This network gives best prediction result of 87.5% on testing data set.

Table 2 Best Neural Network Model for GLCM feature

Model	Multilayer Perceptron with Back Propagation Algorithm
Input Unit	4
Hidden Unit	8
Output Unit	4 (Angry,Happy,Sad,Surprise)
Learning Rate	0.3
Momentum Rate	0.9
Activation Function	Tansig
Stopping Criteria	70 epochs

A Graphical User Interface (GUI) was developed using Graphical User Interface Development Environment (GUIDE) in Matlab. Even though this is not one of this work's objectives, GUI as shown in fig. 14 was developed since it is considered as a good demonstration of completed work.

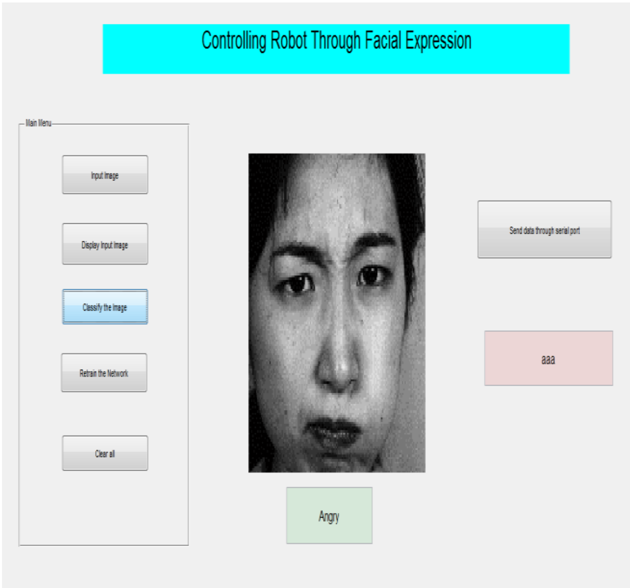


Fig. 14 GUI of facial detection

VIII. CONCLUSION

This research has significance in many fields such as medical, device developer, robotics and Human-Computer Interface researchers. Facial expression also can be utilized to help physically disabled person by creating applications for them. This research uses 120 images from JAFFE database which consist of four facial expressions. These images were preprocessed and feature was extracted using Uniform Local Binary Pattern. The first feature takes advantage of the difference in intensity between the region of the eye and the region of the upper cheek. The second feature calculates the difference between the two eyes and the center bridge of the nose. Back-propagation neural network was used to classify the emotion. The network was able to achieve 87.5% accuracy on testing data. For future work, real-time facial expression detection system recommended. Eyebrow and lip localization based facial expression classification also recommended in order to improve the neural network classification accuracy and reduce the training time since less feature vector would be used.

REFERENCES

1. Morris, C.G and Maisto, A.A. (2002). Psychology: An Introduction. New Jersey: Prentice Hall.
2. Ming-Hsuan Yang; Kriegman, D.J.; Ahuja, N. (2002) "Detecting Faces inImages: A Survey," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol.24, no.1, pp.34-58.
3. Siraj,F., Aziz, A.A,.Sainin, M.S., and Hassin, M. H.M. (2004). The Design of Emotion Detection System to Regulate Human-Agent Interaction.
4. Leon, E., Clarke, G., Sepulveda, F., and Callaghan, V. (2004). Optimised Attribute Selection for Emotion Classification using Physiological Signals. Journal of IEE.
5. Paul Viola and Michael J. Jones, (2004) Robust Real-Time Face Detection, International Journal of Computer Vision 57(2).
6. Mahesh, K.V, (2011). Facial Expression Detection using Principal Component Analysis. Msc Thesis.
7. Ganesh, K.V, Shishir, B, (2008). Recognition of Facial Expression using Gabor Wavelets and Learning Vector Quantization .
8. Shiqing, Z., Xiaoming, Z., and Bicheng, L., (2011) Facial Expression Recognition using Local Fisher Discriminant Analysis.
9. T. Ojala, M. Pietikäinen, D. Harwood, (1996) .A Comparative Study of Texture Measures with Classification Based on Featured Distribution, Pattern Recognition 29 (1) 51–59.
10. Shan, C., Gong, S., and Peter, W.M.,(2008). Facial Expression Recognition Based on Local Binary Patterns: A Comprehensive Study. University of London.
11. JAFFE database URL: http://www.kasrl.org/jaffe.html. Accessed date: 4 January 2012
12. OpenCV. URL: http://opencv.willowgarage.com/wiki/. Accessed date: 8 March 2012.
13. Klimis, S., (2000), Hand Gesture Recognition Using Neural Network , Thesis, School of Electronic and Electrical Engineering, Centre for Vision, Speech and Signal Processing, Surrey University.
14. Sivanandam, S.N., Sumathi, S. and Deepa, S. N., (2011) Introduction to Neural Networks using Matlab 6.0. Mc-Graw Hill
15. Bishop, C.(1995). Neural Networks for Pattern Recognition, Oxford: University Press.
16. Haykin, S. (1999) Neural Networks: A Comprehensive Foundation, 2nd Edition, New Jersey: Prentice Hall.
17. Keally, R. (1999). Artificial Neural Network : An introductory Course. URL: http://www.maths.uwa.edu.au/~rkeally/ann_all/ Accessed date: 18 February 2012.

AUTHORS PROFILE

Mohammed Hussein Baqir, was born in 1956 in Iraq - Baghdad holds a Baklores engineering Electronic and Communications in 1976 and master's degrees in engineering electronic Technology University in 2004-2005, and I am currently a Lecturer at the College of Electrical and Electronic Technique in Baghdad , Crown of scientific research ten search individually as though to me the author of a scientific book titled electrical circuits and measurements , a book is a source of scientific and there is a book in the way of achievement , which holds the title of measuring devices and electronic transformers Medical , which is as a book systematically to the Faculty of electrical engineering and reconciled to God

Mousa Kadhim Wali, received the B.Sc. degree in Electrical Engineering from Baghdad University, Iraq in 1985, the M.Sc. degree in Electronic Engineering from Bradford University, U.K. in 1991, the PhD degree in computer engineering from University Malaysia Perlis (UniMAP). He is currently with the electrical and electronic technical college, Baghdad, as Senior Lecturer. His current fields of interest are in Emotion Recognition, Stress Assessment, Hypovigilance detection, and early prediction of cardio vascular diseases. PID controller.

Majid S. Naghmash, Received His Msc. degree in the satellite communication engineering from University of Technology Baghdad – Iraq-2004 and Doctor of philosophy in wireless and mobile system from University Science Malaysia (USM) Malaysia –Penang -2011. He is currently a head of power engineering department at College of Electrical Power Engineering Techniques-Foundation of technical Education, Ministry of high education and scientific research-Iraq. He has been a full-time lecturer in the Computer Department, Baghdad, Iraq, since March 2012. He also worked as senior researcher in the Iraqi Center of Development and Research since 1994