Hybrid Approach for Facial Expression Recognition using HJDLBP and LBP Histogram in Video Sequences

Mahesh U Nagaral
Dept. Computer Science and Engineering, B.L.D.E.A's V. P. Dr. P. G. Halakatti College of Engineering and Technology, Vijayapur, India.
Email: nagaralm@yahoo.co.in

Dr. T Hanumantha Reddy
Dept. Computer Science and Engineering, RYMEC College of Engineering and Technology, Bellary, India.
Email: thrbly@gmail.com

Abstract—Any kind of compassionate thoughts can't be expressed through words, but it appears on their facial expression. So, the facial expression reveals the emotions of individuals. The recognition of such emotions can be understood correctly or sometimes ambiguously from the opponent. Hence, there is a scope for automatic facial expression recognition (FER) in the context of image processing. The FER system has three different phases: face detection, feature extraction and expression classification. In face detection phase, Viola Jones face detector is used to crop the original image such that only the face region is retained by removing the unwanted region. In feature extraction stage, High-order Joint Derivative Local Binary Pattern (HJDLBP) and Local Binary Pattern (LBP) histogram algorithms are used for extracting features from the cropped image. In last stage, Support Vector machine (SVM) classifier is used in finding the precise facial expression.CK+ dataset has been used for training and testing, which consist of 442 image samples. We have considered six different universal possible expressions such as, happy, anger, disgust, fear, surprised, and sad for identification. The experimental results indicate that the overall accuracy of the proposed system was 74.8%, which is high compare to the results available in literature.

Index Terms—Facial expression, Viola-Jones, LBP histogram, High order joint DLBP, Support Vector Machine.

I. INTRODUCTION

A communication is a process of imparting information or message between two or more human beings. Such a communication is supported either by oral mode or non-verbal mode. The verbal communication mode between two or more persons is through spoken words or hand written letters, and Non- Verbal communication happens through gesture, facial expression, eye contact, etc. As per the history concerned, spoken communication expresses 7% of the message in person communication. Vocal contributes 38% and facial expressions express 55% of the message [16]. The facial expressions are identified by motion to the muscles of the face or the movement of the oral edges, eyebrows; nose called the key points on the face where every expression reviles the inner feelings of the human. An expression on a face is classified into six universal gestures such as happy, angry, surprised, disgust, fear, and sad [7]. Faces and expressions vary from person to person. It's challenging work for Computer system to recognize and classify the expression of different persons. Thus, topic of training in many areas such as psychology, natural sciences, medicine and intelligent human-interactive interface is dealt in automatic Facial Expression Recognition (FER).

An imperative constraint to this FER system is the way that static images catch the peak of the expression, i.e., the moment at which the pointers of feeling are generally marked. In their everyday life, individuals occasionally demonstrate pinnacle of their facial expression amid typical correspondence with their partners, unless for certain cases and for extremely short timeframes.

Another, possibly more essential inspiration FER system is that expression itself is a productive method for correspondence: it's normal, non-meddlesome and shockingly, expression passes on more data than conversations and voice. To create more convenient interfaces that recognize the human computer for expression is required. The automatic FER methodology incorporates:

- Face Detector.
- Facial element extractor.
- Facial expression recognizer

The first component face detector is used to remove the noise and detect the face region of input image sequence using Viola–Jones algorithm. The second compo

nent is about feature extraction, in which facial features are derived based on the facial muscle movements and feature vectors are calculated, here in this system, HJDLBP and LBP histogram feature extraction methods are used. Last component is expression classification; SVM is used for classifying the expressions.

The basic six facial expressions are as follows,

- Anger
- Disgust
- Fear
- Happy
- Sadness
- Surprise

(a) The original image

(b) The cropped image

Anger   Disgust   Fear   Happy   Sadness   Surprise

Fig.1. Six basic expressions from CK+ database [17].

II. RELATED WORK

In this Section we will be discussing briefly about the work carried out by different researchers so far with respect to Facial expression recognition in video sequence.

Behzad Hasani et al [1] have implemented Latent-Dynamic Conditional Random Fields (LDCRFs) and Conditional Random Fields (CRFs) discriminative models for classifying expressions. Active Appearance Model (AAM), Histogram of Oriented Gradients (HOG) and Uniform Local Binary Pattern (U-LBP) are the extraction strategies utilized as a part of work proposed. A set of 68 face landmarks of image are used and those marks are around eyes, nose and mouth. AAM landmarks are used from the CK+ dataset. Principal Component Analysis (PCA) reduces the input vector measurements. To evaluate proposed system CK+ dataset is used. It consists of 593 sequences of videos but only 327 sequences are considered. All expression frames begins with neutral expression and ends with top peak expression. The experiment was performed in two groups, in first group neutral face class is considered and in second group it is excluded. Different methods and models with various hidden sates considered to improve the recognition accuracy. AAM feature gives best result on CRF model, then U-LBP feature is combined with HOG on LDCRF with hidden state variables have shown good result.

Shu An et al [2] new approach utilized to perceive to recognize the facial expression of 3D image. A new method is implemented called Local Threshold Binary Pattern (LTBP) which is related to LBP. The difference of standard gray value of neighboring pixels and central pixel are calculated. Calculated LTBP is fused with Histogram of Oriented Gradient (HOG) to find out multi-feature for 3D expressions. The experiment is done on BU-3DFE database which is the 3D database and the accuracy achieved is 90%. Future research work should focus on few of the expressions types which are difficult to categories such disgust and sad. Need to identify more feature points for the expressions sad and disgust. Very limited numbers of standard datasets are available of 3D images.

Anusha A V et al [3] worked on a novel approach for classification of gender and expression recognition by considering two expressions anger and happy. Geometric and appearance based methods are used in this system. Both gender classification and expression recognitions are done using facial patches. This system works on even with images with low resolution. Haar classifiers used to detect eye and nose area. Detection of lip is implemented by using sobel operator. Quadratic Discriminant Analyzer considered for the predicting gender and also geometrical and appearance feature used for predicting gender. For the better accuracy facial patches gender based classification is used. This system used CK+ and JAFFE databases for testing.

Fuxiang Lu et al [4] author by analyzing the weaknesses of uniform LBP has provided an operator that is Improved Local Boundary Improvement (ILBP) which describes texture of image efficiently. The operators of the ILBP determine an important set of local initials like lines, T-junction and cross-sectional crosses that are ignored by uniform LBP operator. ILBP Operators are available based on a measure to consolidate the model and the number of these LBP codes. The ILBP is unchanged, with respect to changes in gray equality, histograms, unit change operations, and rotation. The main advantage of ILBP compared to the traditional LBP is that it identifies the set of native primitives from non-uniform patterns. Thus, the ILBP function makes perfect between them aboriginal and stability capabilities. Lastly, ILBP PC is attractive and suitable for authentic software because it can be implemented with many operators.

Nan Sheng et al [5] proposed a completely automatic approach for 3D facial expression analysis. Firstly 3D faces are preprocessed to remove spikes and fill any holes if available by cubic interpolation. To get more normalized 3D model resampling and alignment is done. An approach is used to find out 20 landmarks automatically from 3D faces .LBP is applied around the facial key points. Diverse techniques used to distinguish the key points such as nose point, nose width point, external mouth, inward and external eyes. Remote features and LBP features are connected to characterize the 3D facial expression. BU-3DFE is most widely used database to validate the efficiency of FER. It consists of 83 manually identified key points.

Sayed et. al. [6] author has used HJDLBP feature extraction method to find the relationship between multi resolution LBPs. Two datasets JAFFE and YALE used to evaluate the implemented system. In initial step, the face region of each image trimmed utilizing Viola and Jones algorithm. There after the relationships between matching patterns are derived from rectangular divided face image and histograms are calculated for the derived regions. Lastly, the images are trained using SVMs and classified as per the classification class. The comparative result shows the increased efficiency of the system. But the implemented operator requires a lengthy feature vector for histogram.

Jyoti Kumari et al [7] in this work a relative study is done using variety of extraction methods on JAFFE dataset. Automatic FER system support many applications .Mostly widely used FER systems are geometry based and appearance based. Facial Expression recognition support wide areas of application and needs more precise and consistent FER system. This work is presented a study of facial expression recognition. The new technique for extracting these traits is covered. The study is still ongoing to surge the (i) accuracy of predicting the expressions, (ii) applications based on dynamic sequence of videos, (iii) to deal the occlusion.

Yeshudas Muttu et al [8] have utilized Viola - Jones algorithm to the collect the required information from the input image. Face features such as eye, mouth, nose and eye brows are considered for extraction of facial features. Basic LBP algorithm need more time for extracting features, so modified version of LBP algorithm is used which gives better outcome. In preprocessing RGB image is transformed to grayscale image. The image obtained through LBP algorithm is given for histogram plotting. Two types of histograms are plotted one is sparse and another is tight histograms. Sparse histogram is the accurate of image histogram but with this histogram length of the feature vectors will be increased. Whereas tight histogram reduces the feature vector length by reducing the gaps. Later both the histograms are concatenated to get better feature vector. Back propagation algorithm (BPN) classifier output reduces the efficiency so radial basis function neural networks (RBF) classifier used which gave maximum classifications efficiency. Japanese database and Taiwanian database images are used for the comparison of final results.

Dhrubajyoti Boruah et al [9] have employed Hidden Markov Model (HMM) classifier, which gives the good accuracy classification rate. Here they have used algorithm that finds the region of the face automatically and later facial features calculated according to the displacements of inner facial muscle movement. The accuracy of the recognition rate with their own dataset is 82.5% which is less compared to the accuracy rate 90.00% by using MMI database. Another strategy is considered to distinguish the face from the input image sequence and an algorithm implemented to discover face region automatically based on the displacement of the face muscle. Face regions considered are eyes, eye brow and mouth.

But this method will consider only the frontal view of image.

Sayed Mohamad Tabatabaei et al [10] One of the most fascinating ways someone showed their inner feelings through the face expressions. To improve the accuracy of recognition of image expression, here they have used a new version of LB which is called as high order derivative LBP (HJDLBP) and in this work three-dimensional relation between patterns in different resolutions is determined. To calculate the LBP histogram regions radial partitioning method is used and this method increases the classification accuracy. To classify the different expressions, multiple support vector machines are used which are trained effectively. Publicly available dataset JAFFEE is used to evaluate the proposed system.

Hua Gao et al [11] has developed the stress detection system by observing the facial emotions of the driver. The anxiety related emotions considered in this methodology are anger, disgust or blend of both and such feelings are considered as negative feelings. It is imperative to check the condition of being smart and safe to drive enthusiasm and comfort of driving. A system keeps running continuously to permit snappy activity action upon detecting the stress of the driver. A near infrared (NIR) camera placed in car facing towards driver face. In the first step, NIR camera detects and tracks face of the driver then supervised descent method (SDM) is used. This tracker method uses the Viola and Jones algorithm to identify the face region. Face images are normalized using the coordinates of right and left eye. Local descriptors are extracted from the tracked facial landmarks and preserve the facial information. Extracted descriptors are concatenated and PCA applied to reduce the dimensions. SVM classifier is applied to classify the stress emotions.

L. Xie et al [12] initially input video sequence is trimmed and normalized to extricate the face region, at this stage 40 Gabor Magnitude Sequences (GMS) obtained Three orthogonal planes got from LBP(LBP-TOP) are used in extraction of feature to improve progressively. Lastly to find required feature, LBP-TOP histograms of GMSs are used. SVM is a binary classifier, one-to-one SVM classification is used in this work. CK+ database is used to for the experimental purpose, it consists of 210 adults aged, 69% were feminine, 81% were European Americans, 13% were African Americans and 6% other category. In testing phase, second degree polynomial kernel function is utilized after doing analysis with linear and polynomial functions.

S L Happy et al [14] author has implemented real time algorithm of facial expression classification. The analysis of the face is one of the most popular fields Studies in human computer communication (HCI). There are many Implementation in the next-generation user interface, human senses Analyze, behave and behave in a cognitive mode. In this work, algorithm for classification is implemented by using Haar classifiers. The histogram of the different sizes of the block image is a function Vectors and assortment of different appearances using the principal Component Analysis (PCA). This algorithm is

implemented in Determine the exact time of computer expression the complexity of this algorithm is small. Customize the proposed analysis of the appearance is different Condition and intensity display differ from person to person Man. The system uses the image with face to face the grayscale in one person.

Paul Viola et al [15] author proposed a face detection approach which minimizes the calculation time by achieving detection accuracy high. Proposed approach involves construction of face detection structure and proves to be 15 times approximately swift than other previous works. To get true scale invariant, recent face detection approaches work on multiple image scale. The beginning image preprocessing is required to find a person significantly. Advantage of involving integral image is that the face recognition run nearer at the equal time as required to calculate Image pyramid

Mehrabian [16] as author mentioned non-verbal message contains a huge number of symbols and those are very difficult to identify. Silent messages comprises of a detailed inconsistent message about feeling and attitudes. The publication is done in understanding the importance of verbal and non-verbal messages. As per author, three fundamentals version for our liking for the person who expresses oneself with message involving feelings: tone of voice accounts for 38%, words account for 7% and 55% for body language.

III. PROPOSED METHOD

To efficiently recognize the facial expression, the FER system contains of three stages such as face detection, feature extraction and expression classification for efficiently recognize facial expressions. Each step consists of two phases.
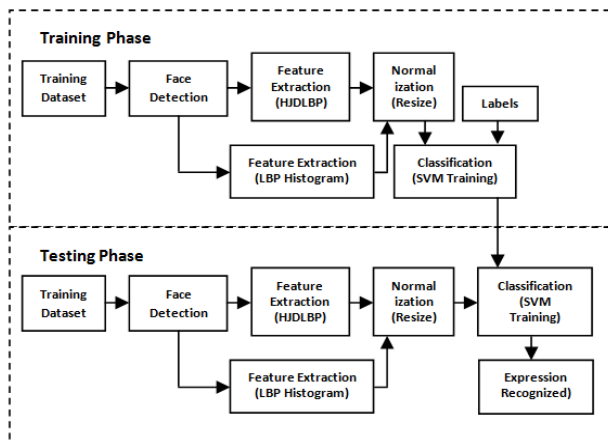

Fig.2. FER system Block Diagram.

FER system block diagram is shown in above Fig. 2. This system is divided into two categories for the classification and recognition of expressions such as:

1. Training Phase: In this phase, the face tracking method is used to detect the phase from the input video sequence. In order to increase the rate of accuracy and computational time resize the detected face. Calculate the feature vector of the input image and label each input image as per the classification.

2. Testing Phase: In this phase, the same face tracking method is used as in training phase to detect the phase from the input video sequence. In order to increase the rate of accuracy and computational time resize the detected face. Calculate the feature vector of the input image and compare the feature vector value with the value obtained in training. If the threshold value of the training features vector matches with testing feature vector return the expression classification otherwise ignore. Sequences of steps carried out by the system are as mentioned below

1. Read the input video sequence/image.
2. Convert the input image to gray scale.
3. Detect the face region
4. Extract feature vector using HJDLBP
5. Extract feature vector using LBP histogram
6. Use SVM to classify the expressions.

3.1 Face Detection

Face region detection is an important phase of any FER system. It has to be accomplished automatically, and true positive rate should be high. This phase is imperative because each image has different sizes and orientation. The most prominent framework currently being used for object detection is known as Viola-Jones [15].

3.1.1 Viola & Jones algorithm

Viola & Jones used an algorithm for detecting the object, which is extensively used in face reorganization. This method gives enhanced results as related with other algorithms, making quick detection and a high true positive rate. Initial phase of the algorithm comprises of extracting the Haar-like features which are also known as rectangular features as illustrated in Fig. 2.1. These rectangular features are applied on 24x24 image of a face (as represented in Fig. 2.2), in different scales and positions.
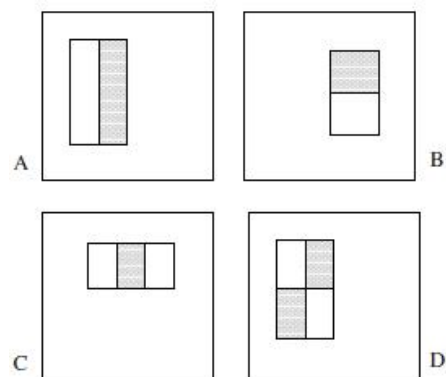

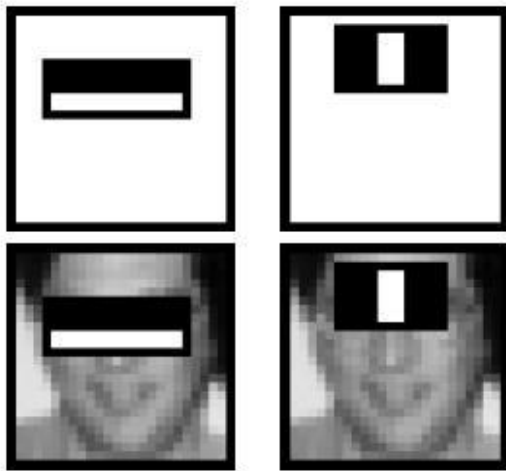Fig. 2.1 Examples of Haar-like features [15]

Fig. 2.2. Haar-like features detected on a face

Since calculating the feature values found to be an expensive procedure, so new concept introduced known as 'Integral Image', it takes constant time for computations. This concept enabled easy and fast way for finding the feature values. However, computing all likely features found to be very expensive. Hence, a new feature selection process was introduced, which uses a modified form of the AdaBoost technique. A strong classifier was created by using this boosting machine learning algorithm.

Finally, cascades of classifiers were used, in which the initial classifiers are simple and from those non-faces are discarded, and stronger classifiers are used to sub-windows the required region. Each stage in cascade classified is trained by using boosting technique. This boosting has an ability to train high efficient classifier by taking the averages considered by the weak classifier.

3.1.2 Detection of face using Viola-Jones

Viola-jones algorithm is highly efficient and universality, so it is used this project to detect and extract the face from the given input video/image. For this, the computer vision system toolbox is used, which is integrated with MATLAB; this tool box contains the vision.CascadeObjectDetector System object which detects objects based on viola-jones algorithms.

The vision.CascadeObjectDetector system object slides the window over the image to detect an object in the input image. Then the cascade classifier is used by the detector to check whether the window comprises the required object. For different scales, the window size varies to detect an object, whereas the aspect ratio retained constant. Then the step(BB = step(face Detector, img)) method is used, which returns the M*4 matrix and accomplishes multi-scale object detection on given input video sequence/image. The returned matrix defines M bounding boxes that contains detected objects. The input to the step method must be grey scale or RGB image.

The left image of Fig. 2.4 show the input image from standard dataset and right side image displays the extracted region by removing non-essential region, and this is accomplished by Viola-Jones detectors.
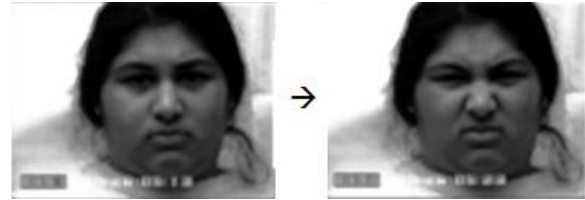
Fig.2.3. Face Detection using viola-jones

Extracted face is converted to double precision and then resized to a 128*128 dimension in 8-bit Gray scale. This resized image is used for feature extraction and classification.

3.2 Feature Extraction

Feature extraction is the significant stage in a FER system. The algorithm chosen at this stage depends on computational properties, type of data used. Hence, in this work HJDLBP and LBP histogram's algorithms are used. In this section will discuss about basic LBP, HJDLBP and LBP histogram.

3.2.1 LBP

Ojala et al. have introduced LBP in the year 2001 [4]. The operator calculates the difference between gray scale intensity of a central pixel c ( ) and the adjacent pixel p ( ) located in radius T. If the computed difference is greater than zero, then "one" should be assigned for that neighbor; otherwise, "zero" will be assigned. As a result, the binary code produced by adjacent P of LBP for the pixel. Eq. (1) defines this operator as,

$$LBP_{R,T} = \sum_{R=0}^{R-1} U(i_q - i_d) \qquad (1)$$

$$U(Z) = \begin{cases} 1 & Z \geq 0 \\ 0 & Z < 0 \end{cases} \qquad (2)$$

Where R is the number of pixels around, and T is the size of radius. The simplest form of LBP is8,1which computes LBP from 8 (R=8) immediately (T=1) neighboring. In multi-resolution operators, LBPs with various R and Tlike LBP8,1, LBP16,2and LBP24,3are calculated. Although these LBPs describe these distinctly differently, the relationship between these operators can be used to describe the texture.

3.2.2 HJDLBP

HJDLBP has been inspired by describing the many new textures proposed by Dan et al. [4].In fact, there are some relationships between LBPs on different scales. For example, in condition that LBP $_{8,1}$ is equivalent to 4 for a pixel, which essentially translates an edge, then there is high probability that $LBP_{16,2}$ of that pixel equals to 8 which also represents an edge. Therefore, the representation (4,8) signifies an edge structure. Dan et al. presented this relationship for the first-order DLBP which is named as a joint local binary pattern (JLBP).In HJDLBP operator, the relationships between high-order DLBPs are considered for texture classification. The relationships between these signal properties would definitely be an efficient image texture descriptor. Eq.(3) formally defines

the relationship between various orders of DLBP including original LBP.

$$HJDLBP_{s1,s2}^{msiv2}(g_c) =$$
$$horzcat\left(LBP_{s1}^{msiv2}(g_c), LBP_{s2}^{msiv2}(g_c)\right) \quad (3)$$

in Eq.(4.3), s1and s2 are two particular (Q, S), m represents the order of LBP, side note rotation invariant pattern and denotes uniform pattern and term horzcat denotes the horizontal concatenation. The above equation can be extended to include more resolutions for numerous orders of DLBP. Fig. 2.4 and Fig. 2.5 illustrate examples for calculating.

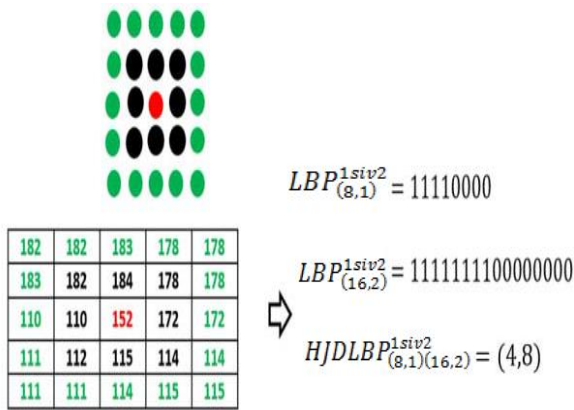$HJDLBP_{(8,1)(16,2)}^{1siv2}$ and $HJDLBP_{(8,1)(16,2)}^{2siv2}$ respectively.

$$LBP_{(8,1)}^{1siv2} = 11110000$$
$$LBP_{(16,2)}^{1siv2} = 1111111100000000$$
$$HJDLBP_{(8,1)(16,2)}^{1siv2} = (4,8)$$

Fig.2.4 An Example of Computing $HJDLBP_{(8,1)(16,2)}^{1siv2}$[9]

$$LBP_{(8,1)}^{2siv2} = 00001100$$
$$LBP_{(16,2)}^{2siv2} = 00001111$$
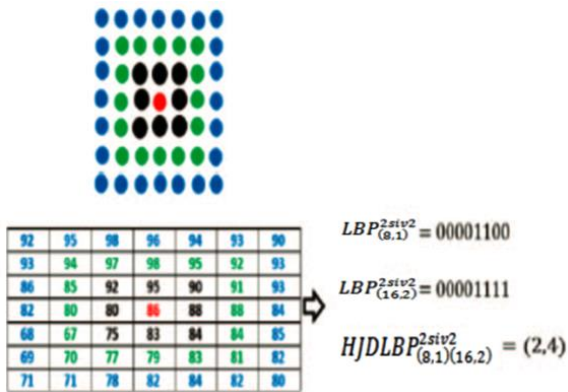$$HJDLBP_{(8,1)(16,2)}^{2siv2} = (2,4)$$

Fig.2.5. An Example of Computing $HJDLBP_{(8,1)(16,2)}^{2siv2}$[9]

3.2.3 Constructing the feature vector with HJDLBP and LBP

HJDLBP and LBP histograms are used in order to extract meaningful information from the faces.At first step, the face detected through viola-Jones algorithm is resized to 128*128 pixels, this will help in reducing the computational time and increases the accuracy rate. Feature vector of the input image is calculated by using HLDLBP then that feature will be represented in the form of the histogram graph as shown in Fig. 2.6.

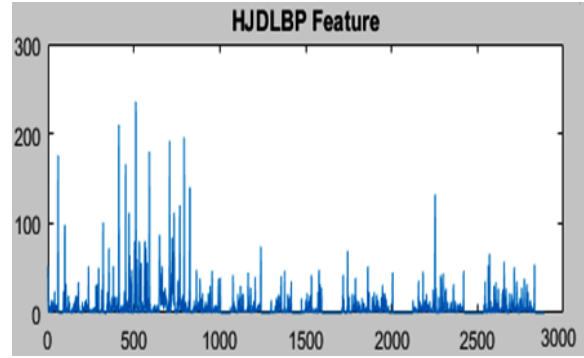Fig.2.6 HJDLBP Feature

Feature vector obtained is resized to 1*332, which will convert the matrix for to feature vector to array format. This is done because SVM has to be trained in array format.
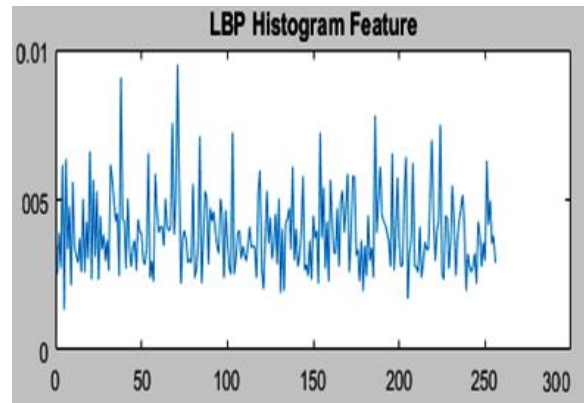

Fig.2.7 LBP Feature Histogram

In final step, LBP histogram of the feature vector is calculated using the hist function as shown in above Fig. 2.7 and then each trained data is labeled according to expression class it belongs to.

3.3 Classification

The last stage of FER system comprises of a model that is trained to accomplish expression classification on new input images. These model uses, machine learning classifier called as Support Vector Machines (SVMs), which accepts input from the feature extraction module, the feature vectors, and learns the patterns that distinguish one expression from the other. Below first subsection introduces the concept of SVM and later part gives detailed explanation of how SVM is used in the FER system.

3.3.1 Support vector machines

Support Vector Machines are most widely used methods for pattern classification. It is a supervised learning model, and uses labeled examples during training process. The trained examples are analyzed and then derive the boundary that will linearly separate the samples of the corresponding classes. This method finds the best boundary by maximizing the margin. Separating data with SVM's optimal boundary is shown in Fig. 2.8
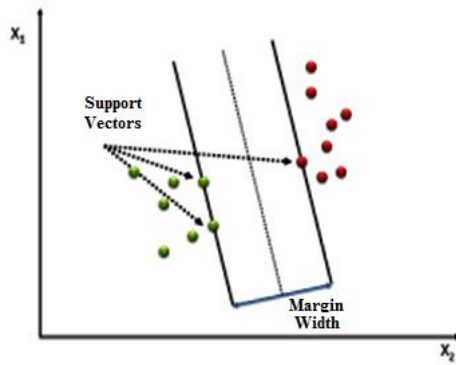
Fig.2.8 Separation of classes [6]

As illustrated in Fig. 2.8 the linear classifier which separates samples is represented in the mathematical form: (equation is taken from [13])

$$f(y) = W^s Y + c \qquad (4)$$

Where W denotes weight vector, c is bias and Y is vector training examples, corresponding to the classes z= {1,-1}. The objective is finding the best values of W and c, which corresponds to the maximum boundary, each training example yi can be described as (equations taken from [13])

$$y_{i \cdot W + c} \geq +1 \ \ if \ z_i = +1 \qquad (5)$$

$$y_{i \cdot W + c} \leq -1 \ \ if \ z_i = -1 \qquad (6)$$

SVM uses complicated functions called 'kernels', these functions map the data points into a higher-dimensional space and hyper plane would be able to separate the samples. The selected kernel function can be: a polynomial kernel or a radial basis function.

3.3.2 Using SVM to classify the expressions

SVM is chosen as a prediction model in FER system, because it is a composite robust model and over-fitting is very minimal. In feature extraction stage, set of feature vectors are produced for the images, and those are used as training examples. These feature vectors are normalized and are fed into the linear SVM model. By using SVM, we can classify data into two different groups. Therefore, there are two strategies for multicasts classification with SVM, first one is one-versus-all method and second one are one-versus-one method. In our work, one-versus-one method is used, in which number of required SVMs for classifying n groups equals to k(k-1)/2. Steps carried out during SVM training and testing in FER system are as follows.

Step 1: Feature vectors of the training images obtained in feature extraction stage are normalized.
Step 2: SVM kernel function is created.
Step 3: SVM training kernel function is created using SVM kernel function.

Step 4: SVM training is performed using function svmtrain().
Step 5: Save SVM train model
Step 6: Save SVM trained data
Step 7:Calculate feature vectors of the testing images.
Step 8: SVM testing kernel is created using SVM kernel function
Step 9: Compare feature vector values with the SVM trained data.
Step 10: classify the image to the expression class it belongs.

IV. RESULTS AND DISCUSSIONS

The efficiency of the proposed system is evaluated by performing three distinct experimentations on CK+ dataset. First as shown in Table 1 complete 30 images are trained of distinctive samples and 70 images are tested, the overall efficiency for this condition is 76%. Next as shown in Table 2 total 40 images are trained of different samples and 60 images are tested, the overall efficiency for this condition is 77%. The randomly chosen images of each expression from total 442 images for training purpose and remaining images are used for testing in third experimentation. Confusion matrix for the same has been shown in table 3. The overall accuracy rate(Performance) of the system is 74.8%.The results of the confusion matrix gives proof to conclude that the FER system correctly classifies all expressions such as happy, anger, surprise, sad, fear, disgust in optimal as shown in Fig. 4.1.Fig.4.2 and Fig. 4.3.

Table 1. Confusion Matrix of Emotion Detection for 30 Training Samples and 70 Testing Samples

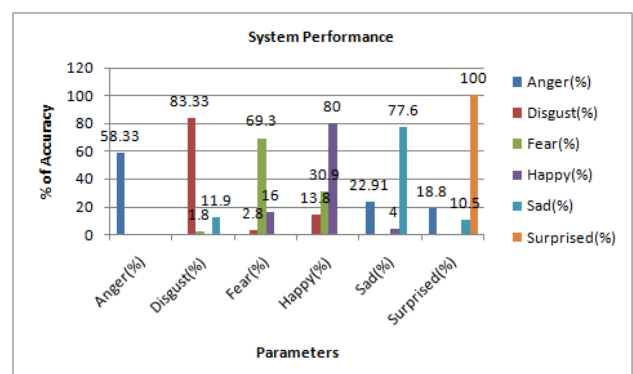| Facial Expressions (%)	Anger	Disgust	Fear	Happy	Sad	Surprised
Anger	58.33	0	0	0	22.91	18.80
Disgust	0	83.33	2.8	13.8	0	0
Fear	0	1.8	69.3	30.9	0	0
Happy	0	0	16	80	4	0
Sad	0	11.9	0	0	77.6	10.5
Surprised	0	0	0	0	0	100

Fig.4.1 System Performance

Table 2. Confusion Matrix of Emotion Detection for 40 Training Samples and 60 Testing Samples

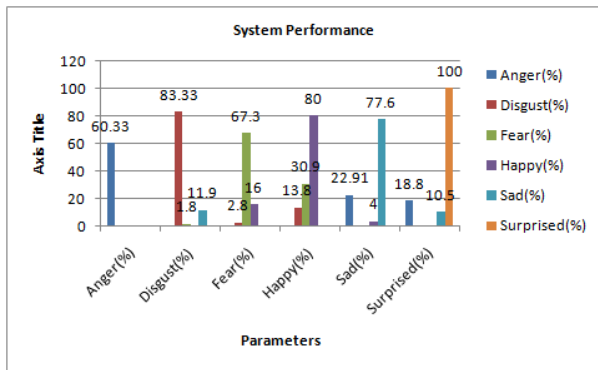| Facial Expressions (%)	Anger	Disgust	Fear	Happy	Sad	Surprised
Anger	60.33	0	0	0	22.91	18.80
Disgust	0	83.33	2.8	13.8	0	0
Fear	0	1.8	67.3	30.9	0	0
Happy	0	0	16	80	4	0
Sad	0	11.9	0	0	77.6	10.5
Surprised	0	0	0	0	0	100

Fig.4.2 Overall System Performance

Table 3. Confusion Matrix of Emotion Detection for whole CK+ Dataset

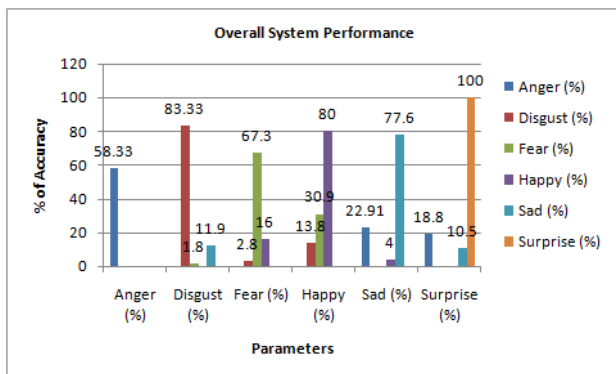| Facial Expressions (%)	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	58.33	0	0	0	22.91	18.80
Disgust	0	83.33	2.8	13.8	0	0
Fear	0	1.8	67.3	30.9	0	0
Happy	0	0	16	80	4	0
Sad	0	11.9	0	0	77.6	10.5
Surprise	0	0	0	0	0	100

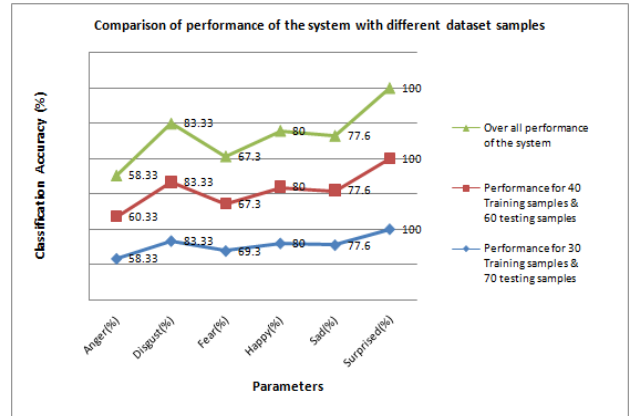Fig.4.3 Overall System Performance

Fig.4.4 Comparison of classification accuracy of the system with different dataset samples

V. CONCLUSION

As automatic facial expression recognition has many applications in the field of human-computer interaction (HCI), there has been an abundant importance enhancing existing techniques as of late. However, there are yet many difficulties and issues to resolve in such systems, particularly in the field of performance and accuracy improvement. Here we have employed high-order derivatives of LBP in conjunction with LBP to improve accuracy. In order to carry out the experiments, initially face region from the image is extracted using Viola-Jones algorithm. After that, features vectors are calculated from face image blocks using HJDLBP and LBP histogram. Lastly, multiple SVMs are trained based on the extracted features to classify images and utilized for facial expression classification using CK+ database.

REFERENCES

[1] B. Hasani, M. M. Arzani, M. Fathy, and K. Raahemifar, "Facial expression recognition with discriminatory graphical models," *2nd International Conference of Signal Processing and Intelligent Systems,* 2016, pp. 1–7.

[2] Shu An and Qiuqi Ruan, "3D Facial expression recognition Algorithm using Local Threshold Binary Pattern and Histogram of Oriented Gradient", *IEEE 13th International Conference on Signal Processing (ICSP)*, 2016, November, pp. 265-270.

[3] A. V. Anusha, J. K. Jayasree, A. Bhaskar, and R. P. Aneesh, "Facial expression recognition and gender classification using facial patches," *International Conference on Communication Systems and Networks*, 2016, pp. 200–204.

[4] F. Lu and J. Huang, "An improved local binary pattern operator for texture classification," *IEEE International Conference on Acoustics, Speech and Signal Processing* ,2016, pp. 1308–1311.

[5] N. Sheng, Y. Cai, C. Zhan, Changyan Qiu, Yize Cui, and Xurong Gao, "3D facial expression recognition using distance features and LBP features based on automatically detected keypoints," *9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI),* 2016, pp. 396–401.

[6]   S. M. Tabatabaei and A. Chalechale, "HJDLBP: A novel texture descriptor and its application in face recognition,"*The International Symposium on Artificial Intelligence and Signal Processing (AISP),*2015, pp. 154–158.

[7]   J. Kumari, R. Rajesh, and K. M. Pooja, "Facial Expression Recognition: A Survey," *Procedia Computer Science*, vol. 58, pp. 486–491, 2015.

[8]   Y. Muttu and H. G. Virani, "Effective face detection, feature extraction & neural network based approaches for facial expression recognition," *International x Conference on Information Processing*, 2015, pp. 102–107.

[9]   D. Bforuah, K. K. Sarma, and A. K. Talukdar, "Different face regions detection based facial expression recognition," *2nd International Conference on Signal Processing and Integrated Networks,* 2015, pp. 459–464.

[10]   S. M. Tabatabaei and A. Chalechale, "Combination of high order joint derivative local binary pattern (HJDLBP) and radial partitioning for automatic facial expression recognition," *International Congress on Technology, Communication and Knowledge ,2015, pp. 86–91.*

[11]   H. Gao, A. Yuce, and J.-P. Thiran, "Detecting emotional stress from facial expressions for driving safety," IEEE International Conference on Image Processing, 2014, pp. 5961–5965.

[12]   L. Xie, H. Wei, W. Yang, and K. Zhang, "Video-based facial expression recognition using histogram sequence of local Gabor binary patterns from three orthogonal planes," *Proceedings of the 33rd Chinese Control Conference,*2014, pp. 4772–4776.

[13]   Andrew R. Webb and Keith D. Copsey. Statistical Pattern Recognition. Wiley, 2011.

[14]   I. Kotsia and I. Pitas, "Real time facial expression recognition from image sequences using support vector machines," *IEEE International Conference on Image Processing,* 2005, p. II-966.

[15]   P. Viola and M. J. Jones, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, May 2004.

[16]   A. Mehrabian,"Communication without Words", *Psychology Today*, Vol. 2, No. 4, pp. 53-56, 1968.

[17]   P. Lucey, J.F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and R. Institute, The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression, In IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2010: 94–101.

Authors' Profiles

Mahesh U. Nagaral, male, is received B. E. (Electrical and Electronics Engineering) from BLDEA's V.P Dr.P.G Halakatti College of Engineering & Technolgy,Vijayapur under Karnataka University, Dharwad and Master of Technology (Computer Science and Engineering) from P. D. A. Engineering College, Gulbarga under Visvesvaraya Technological University, Belagavi. He is currently a research scholar registered to V.T.U., Belagavi in the area of Image and Video processing working on the issues of Facial Expression Recognition from Video. Currently, he is working as Assistant Professor in Department of Computer Science and Engineering, B.L.D.E.A's V. P. Dr. P. G. Halakatti College of Engineering and Technology, Vijayapur. His subjects of interests are Image and Video Processing, Software Engineering, Programming Languages, Data Structures and Object Oriented Programming. He has total of 03 publications in the area of Image and Video Processing.

Dr. T. Hanumantha Reddy, male, is received B.E. (Electronics and Communication) from Gulbarga University, Gulbarga and Master of Engineering in Computer Engineering from Mysore University, Mysore. He has completed his Ph.D. from J. N. T. U., Hyderabad in the area of Image Processing. Currently he is working as Professor and Head in the Department of Computer science and Engineering, R. Y. M. E. C., Bellary. His subjects of interest are Image and Video Processing, Data Communications and Computer Networks, Computer Architectures and Multimedia computation and Communication. He has total of 26 publications in the area of Image Processing.