



## A Salient Binary Coding Scheme for Face and Expression Recognition from Facial Images

Sammaiah Seelothu<sup>1\*</sup>      Konakanchi Venugopal Rao<sup>2</sup>

<sup>1</sup>*Department of Computer Science Engineering, Jawaharlal Nehru Technological University,  
Hyderabad, Telangana, India*

<sup>2</sup>*Department of Computer Science Engineering, G Narayanamma Institute of Technology and Science,  
Hyderabad, Telangana, India*

\* Corresponding author's Email: sammaiahmanuu@gmail.com

---

**Abstract:** The automatic Face Expression Recognition is very important in computer vision since it has widespread applications in real time. However, the main issue is the design of an efficient descriptor that describes the appearance changes on the face. Towards such design, in this paper we have introduced a novel binary compact coding called as Edge Adaptive Local Directional Binary Pattern (EALDBP) that encodes the texture features of facial image. The proposed EALDBP is resilient to noise and illumination variations as well as it encodes the directionality of expressions. The directionality is discovered by the computation of edge responses through different compass masks. Further to encode intensity variations, we have considered the relationship between gradient pixels in the local neighborhood. Finally, a compact face descriptor is constructed with the help of Histograms computation over the encoded face image. To evaluate the performance, we tested our approach on two datasets; they are CK+ and JAFFE. The average recognition accuracy of these two datasets is observed as 90.3898% and 93.5922% for JAFFE and CK+ respectively. These accuracies demonstrate that the proposed approach outperforms the several state-of-art methods.

**Keywords:** Face expression recognition, Face recognition, Directional encoding, Magnitude strength encoding, Directional binary pattern, Image descriptor, Edge response.

---

### 1. Introduction

In the last few decades, a lot of research has been carried out over Face Expression Recognition (FER) in the field of computer vision [1]. Generally human beings employ different cues to express their emotions such as voice, hand gestures and face expressions. Among these cues, the face expressions have major contribution and it is approximately 55% of human communications. Hence, a wide spread utilization of FER is employed in several applications related to operator fatigue detection, automatic tutoring systems, music for mood, lie detection, face expression synthesis, automatic counseling systems, security etc. [2]. For example, the robots can easily interact with human beings if they can analyze the facial expressions thereby they can discover the emotional states. In this manner, they can be

employed in healthcare systems for the detection of human's mental states such that their life quality is improved. Generally, there are two kinds of expressions; they are positive expressions and negative expressions. Positive expressions include Pleasure and Happy (demonstrates that the human being is healthy) and negative expressions include sad and anger (demonstrates that the human being is unhealthy). Thus an effective FER can help in the improvisation of mental emotion state by analyzing the behavior patterns [3, 4].

The major components of a typical FER system are Face Image representation and Classification. The face image representation is a most important aspect in which the face image is represented or described in such a way that the FER is robust to all possible facial variations like noise, pose and illuminations. A FER is more effective if the face descriptor is able to lessen

the intra-class and maximize inter-class variations. A good face descriptor can assure an accurate and efficient recognition process. So many earlier studies are focused in this direction and they are broadly classified into two classes [5-7]; they are geometric based approaches and appearance based approaches. The geometric based approaches [8-10] describe a face image by the computation of geometric relationships between the positions of several facial elements like nose, ear, eye and mouth. Generally, these approaches represent a face image through its position, shapes, and angles between face elements. However, the major problem is that they require reliable and accurate facial feature detection and tracking which is a complex issue in many real time applications. Next, the appearance based methods [11] describes a face image through its texture. To extract the texture, they employ filters either on the local regions of facial image or on the whole facial image. Compared to geometric based approaches the appearance based approaches have achieved higher recognition rate but they also have several limitations. The first limitation is loss of information and second limitation is limited performance in unconstrained environments.

To solve these problems, we have proposed a new face descriptor, called as Edge Adaptive Local Directional Binary Pattern (EALDBP) that encodes the structural information as well as intensity variational information. According to EALDBP, we represent every pixel of face image with an eight bit binary code and it is derived based on the directional and intensity information. To do this task, we need edge responses of the face image and they are obtained through two different compass masks in eight directions; they are Robinson compass mask and Gaussian compass mask. Then from all the directions, we search for top positive and negative directions to produce an effective descriptor which can describe different textures with similar structural patterns. Further to encode intensity variations, we have encoded the relationship of neighbor pixels with center pixel. Since our method used more pixels into consideration, the encoded pixel conveys more information. Finally, we compute histograms of the encoded face image to get final face descriptor.

Rest structure of the paper is organized as follows: in section 2 we discuss the details of literature survey. In section 3, we explain the complete details of proposed face expression recognition system. Section 4 explores the details of experimental results and performance analysis. Finally, we present the concluding remarks in section 5.

## 2. Literature survey

In the literature survey, there are many appearance based methods for holistic representation and local representation. The holistic representation methods were applied on the whole image while the local representation methods applied on the local regions of image. Eigen faces [12] and Fisher faces [13] are two best examples for holistic representation and they have been widely used in face expression recognition [14]. However, these methods are not robust for pose variations and illumination changes in the facial images. Due to several variations in the facial expression, the simple Eigen values want have significant effect on the expression recognition. Moreover, the holistic representation misses some important information about minor movements in face.

On the other hand, the local representation methods are robust to the pose and illumination changes because they convey local information of a pixel with respect to its neighbor pixels which have strong correlations. Histogram of Gradients (HoGs) and Local Binary Pattern (LBP) [11] are the two popular and base methods for local representations. J. Kulandai Josephine Julina, T. Shree Sharmila [15] used HoG [21] for face representation. According to this method, initially the facial image is divided into small regions and then the gradient histograms are measured for every region. However, they didn't focus on the statistical relationship between adjacent pixels which are strongly required for facial expression recognition. To solve this problem, LBP is introduced and Bhupendra Singh et al. [11] used LBP for facial expression recognition. In the LBP, a pixel is represented with a binary code and it is derived based on the relationship with neighbor pixels in a local region. Compared to HoG, LBP achieved a superior performance in expression recognition. Hence most of the current research is focused on LBP and developed several variants of LBP. N. T. Cao et al. [16] proposed a FER system based on LBP and Support Vector machine (SVM). Before LBP, this method adopted a preprocessing step to divide the face image into non-overlapping square regions. N. Kauser and J. Sharma [17] initially segmented the facial elements like mouth, nose, and eyes and then applied LBP to encode each pixel of all these elements. For every element a template is created and it was trained and classified through Neural Network. However, the major problem with LBP is loss of information. During the computation of an LBP of a pixel, it considers neighbor pixels intensity and subtracts from center pixel and the encoding is done based on the subtracted values. The

subtraction process results in loss of information. Moreover, the LBP based face expression recognition system is much sensitive to noise.

To overcome such problems, several variants of LBP are developed. Recently, M. Goyani and N. Patel [22] developed Local Mean Binary Pattern (LMBP) that encodes the texture and global shape of face. LMBP of a pixel is measured based on the weighted mean of neighbor pixel's threshold intensity values. Instead of subtracting the neighbor pixel intensities directly from center pixel intensity; they are subtracted from mean of patch. This process creates an additional loss of information compared with normal LBP. A weight matrix has defined which is an addition process but not required.

T. Jabid et al. [18] proposed a new variant of LBP called as Local Directional Pattern (LDP) to represent the face expressions. In LDP, a pixel is encoded according to the relative magnitude strength of edge response values in eight directions. For the computation for edge responses of facial image at eight directions, kirsch compass mask is employed. In this method, every pixel is encoded through eight binary code words which create a more confusion for recognition system.

Next, T. Xu et al. [19] studied the effect of weightings on the modular's LDP on the recognition rate. This approach adopted the combination of Principal Component Analysis (PCA) and LDP subspace for feature count reduction. Nearest Neighbor Classifier is employed for classification. However, the PCA assisted feature count reduction has less impact on the variations of expression in a single facial image.

Next, Md Hasanul Kabir et al. [20] proposed a new variant of LDP called as Local Directional Pattern Variance (LDPV) that encodes the contrast as well as texture information of facial image. The LDPV is a collection of LDPs weighted by their respective variances. PCA is employed for reducing the dimensions of features and SVM is adopted for classification. Encoding of contrast information with the help of local variance is an additional benefit of this method because it makes the system robust to illumination variations. However, local variance has only less impact on the accurate recognition because some expression looks like an *d* at such conditions, the recognition of similar expressions like *sad* and *fear* becomes complex.

Local Derivative Patterns (LD<sub>e</sub>P) [23] is one more variant of LDP and it is a higher-order local pattern descriptor, employed for face recognition. LD<sub>e</sub>P encodes the directional pattern features based on local derivative variations. Unlike the LBP process which encodes the relation between center

pixel and neighbor pixels, the LD<sub>e</sub>P considers the higher order derivatives and encodes the pixel based on the neighborhood relations. Although these methods use more information to encode the pixel which results in a stabilized binary code, they still encode the information in the same manner of LBP. Despite this simple coding strategy, these methods also discard most of the information from the neighborhood. Moreover, the local derivative patterns have good performance only if there are sharp edge variations. For smooth variations, derivative patterns contribute only less contribution. Local Directional Number Pattern (LDN) [24] is another variant that encodes structure of local neighborhood based on the directional information. In LDN, a pixel is encoded according to the relative directional strength of edge response values in eight directions. Kirsch and Gaussian compass masks are employed in eight directions to obtain edge responses. However, LDN considered only directional information, means only two neighbor pixels information is encoded which proves that the information loss is high.

Some more feature extraction methods are developed based on LBP combining it with several handcrafted methods. Combining the covariance matrix and K-L transform with Extended LBP (ELBP), Guo et al. [25] proposed a FER recognition system. First, ELBP is used to extract the feature matrix of facial expression images and then covariance matrix transform is accomplished for dimensionality reduction. And finally recognition is done using SVM classifier. This approach is not considered the directional information which has major role in the improvisation of recognition performance. KLT ensures reduced dimensions but discards much information.

Abdul, M., and Holambe, R. S., [26] combined Directional wavelet transform (DIWT) with LBP to extract the facial features. Initially the facial image is decomposed into directional sub bands and an adaptive direction selection method is accomplished based on quad tree partitioning to obtain top-level DIWT sub bands. Next the LBP histograms are extracted from the selected top bands to obtain local descriptive feature set. The wavelet decomposition of image ensures robustness against multi-resolution property but not against illumination, and poses variations. Moreover, the directionality is also not considered by which the expressions like *sad* and *fear* may get misclassified.

Qian Zhang et al. [27] combined the LBP with 2-D Gabor wavelet filter [29] for face expression recognition. The LBP ensure invariance against illumination while the Gabor wavelet filter ensures

invariance against pose and expression variations. But the necessity of directional information is there in expression recognition which was not considered here. Moreover, the Gabor wavelet transform breaks the relationship between neighbor pixels consequences to the edge or boundary discontinuities.

The FER system proposed by I. M. Revina and W. R. S Emmanuel [28] developed a new descriptor called as Dominant Gradient Local Ternary Pattern (DGLTP). EDLTP is an extension to the Local Ternary Pattern (LTP) [30] and used to extract the local dominant texture features of a facial image. At preprocessing stage, to remove the noise, Enhanced Modified Decision Based Unsymmetric Trimmed Median Filter (EMDBUTMF) is employed. After preprocessing, the histogram features are extracted through DGLTP and then fed to SVM classifier for expression classification. However, the LTP need additional bits for pixel representation. Furthermore, the negligence of directional had shown significant impact on the recognition performance.

Kamal A. El Dahshan et al. [41] considered the Deep Belief network for emotion classification from facial image. However, this method not robust for illumination, pose and noise variations because there is no method for feature extraction they have used a classifier directly which creates a huge computational burden over the system. Our main intention is to gain an improved recognition with less computational; burden. This is possible only when feature count is less followed by most discriminative features.

### 3. Proposed approach

### 3.1 Overview

In this section, we explain the complete details of proposed face ad face expression recognition system. The overall schematic of proposed system is depicted in Fig. 1. According to this figure, we have proposed a new face descriptor, Edge Adaptive Local Directional Binary Pattern (EALDBP) Histogram (EH) that encodes the intensity variations and structural variations of a facial image. The proposed EALDBP is an eight bit binary code that was derived based on the relation of a pixel with its neighbor pixels. Under this evaluation, we have used two different types of compass masks to measure the strong edge responses of a face image. They are (1) Robinson Compass mask and (2) Gaussian Compass mask.

The Robinson compass mask alleviates the edge responses of a facial image. Next, the Gaussian compass mask alleviates the illumination variations and derives strong edge responses which are resilient to noise. For each mask, the edge responses are derived in eight directions. Then from all directions we select top positive and top negative directions and each direction is represented with three bits. Along with these directions, we also search for second top positive and negative magnitudes and each magnitude is represented with one bit. Hence each pixel is represented with an eight bit binary code. Since we considered the directions of edge responses, the obtained face descriptor is robust for different facial images with similar structural patterns. Further, our binary code also encodes the information of neighbor pixels, hence it conveys more information.

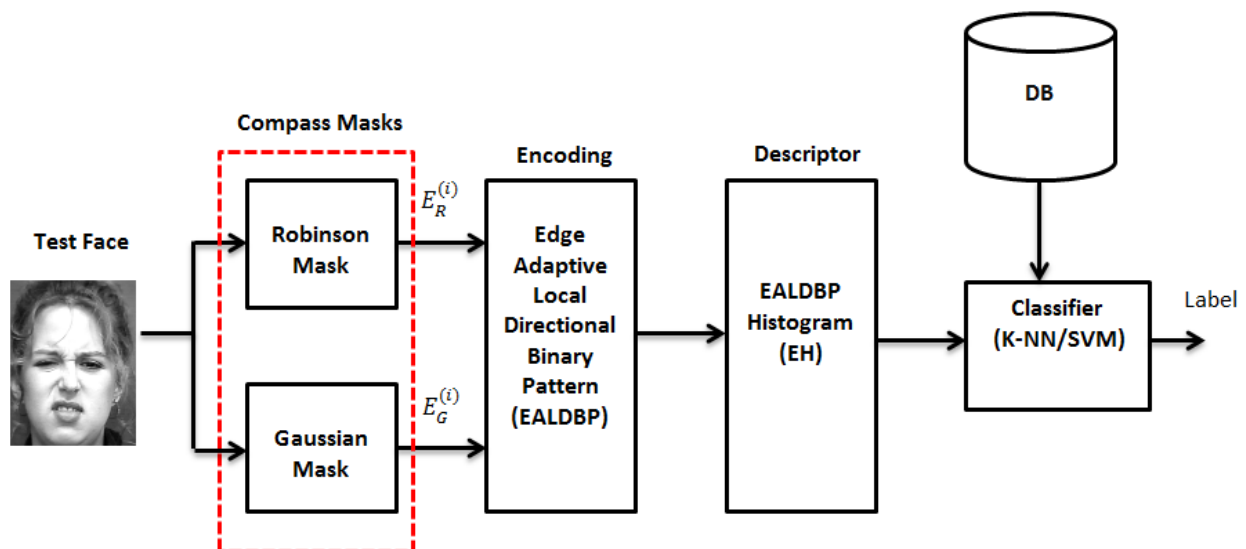


Figure. 1 Block diagram of proposed system

Finally, for the obtained EALDBP at each mask, we compute Histogram thereby the face is described with EH. The important point is that we need an exact face region and for that purpose we employ the Viola-jones algorithm [31].

### 3.2 Compass mask

To represent a face image with EALDBP, we need edge responses and they are derived through compass masks. In this method, we used two compass masks to get the edge responses; they are 1) Robinson Compass mask [32] and (2) Gaussian Compass mask [33]. Before subjecting the face image to get the edge responses by the convolution of masks, initially it was transformed into gradient domain from pixel domain [35]. In this approach, we used gradient space instead of pixel space means for a given facial image, we derived gradients and they are used for further processes. The main reason behind the transformation is that the pixel space ignores the relationship between adjacent pixels thereby it never conveys more information about the facial variations. Unlike the pixel space, the gradient space considers the relation between neighbor pixels thereby more information is available. The relation between neighbor pixels in gradient space is more advantageous because it reveals the structural information of face image. Moreover, the gradient space has more discriminative capability and it can help in the recognition of different faces as well as expressions. Both the compass masks are applied over the gradient space of image such that the main underlying structure of facial image is obtained.

#### 3.2.1. Robinson compass mask

In the facial expression recognition system, the shape of facial image changes according to the facial expressions. If the proposed face descriptor is able to capture such changes, then the recognition system becomes more robust for facial expressions. Moreover, the change features are limited in number thereby the computation burden is also reduced. Hence such kind of features extraction has prime importance in face or expression recognition. Further the boundaries of facial features have higher magnitudes when compared with the magnitudes of smooth regions. Hence we employed the most popular Robinson edge operator to extract the boundaries of face images as edge responses. Robinson compass mask is also termed as direction mask. In this operator, we take one mask and it is rotated in eight directions namely, North (N), North West (NW), West (W), West South (WS), South (S), South East (SE), East (E), East North (EN). This

mask is a symmetric mask hence it generates the same magnitude responses in opposite directions with different signs. For example, the magnitudes of edge responses of second and sixth masks are same with different signs. For example, consider the magnitude of a pixel in the second edge response is 456, and then the magnitude of same pixel in sixth edge response is -456. Hence out of eight edge responses, we consider only four edge responses. The obtained masks in eight directions are shown in Fig. 2 and their Edge responses over a face image are shown in Fig. 3.

#### 3.2.2. Gaussian compass mask

Gaussian compass mask is employed over the face image to make the recognition system robust to noise and illumination variations. Also, the Gaussian compass mask smoothens the face image which makes the computation of gradient space more stable. Unlike the Robinson compass mask, the Gaussian compass mask is asymmetric and it was generally used to derive the edge responses over smooth regions of faces.

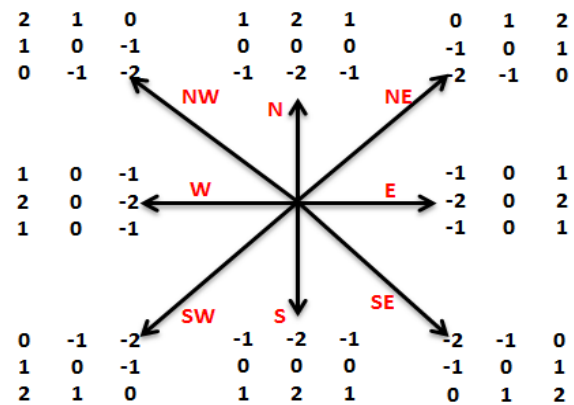


Figure. 2 Robinson compass masks in eight directions

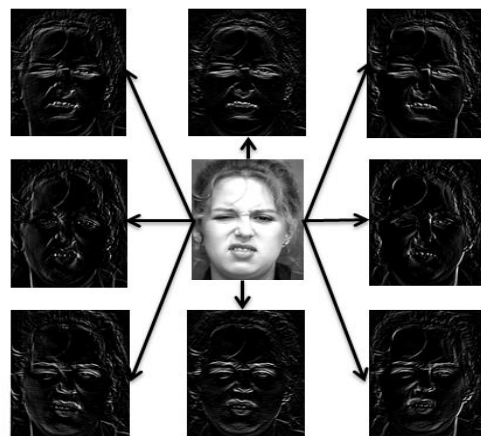


Figure. 3 Edge responses of Robinson compass masks in eight directions

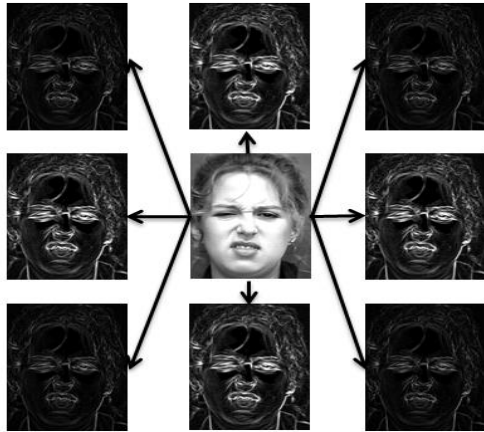


Figure. 4 Edge responses of Gaussian masks in eight directions

The Gaussian mask is computed according to the following formula;

$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (1)$$

Where  $(x, y)$  is the position and  $\sigma$  is the standard deviation of compass mask.

Inspired with the Robinson compass mask, we generate eight masks for eight directions by rotating  $G_{\sigma}$  in eight directions with an angle of  $45^{\circ}$ . Means the angular deviation between successive compass masks is  $45^{\circ}$ . The edge responses obtained after the convolution of Gaussian masks with a face image in eight directions is show in Fig. 4.

### 3.3 Encoding

In the proposed encoding scheme, we produce an EALDBP for each pixel in the facial image. The EALDBP is an eight bit binary code. In this code, the first three bits (1, 2, and 3) are belongs the direction of first top positive edge response. The next bit (i.e., 4<sup>th</sup> bit) belongs to the magnitude deviation of center pixel and its neighbor pixel which has is second top positive edge response. The next three bits (i.e., 5, 6, and 7) are belongs to the direction of first top negative edge response. The final bit (i.e., eight bit) is belongs to the magnitude deviation of center pixel and its neighbor pixel which has is second top negative edge response. In the obtained binary code of a pixel, we can get more information because we have encoded the formation of all pixels in the neighborhood. However, in the LDN, only directional information is encoded such that it has too much information loss.

The major properties of EALDBP are; (1) Gradient magnitude and direction are used for superior representation of shape of the expression related features. (2) Gaussian Compass mask is

resilient to noise and illumination variations in facial image. (3) Robinson compass mask is much efficient and computational ensure a lessen burden due to its symmetric property. (4) Directional information along with magnitude deviations information lessens the information loss. In the most popular LBP, only magnitude deviations are considered to encode the pixel and in the LDN only directional information is used to encode the pixel. Unlike these two methods, we have considered both gradient directions as well as magnitude deviations thereby the obtained EALDBP convey more information about the structure of facial image. The encoding scheme is applied over the edge responses obtained through both the compass masks and they are illustrated in the following sub-sections;

#### 3.3.1. Robinson encoding ( $EALDBP^R$ )

In this phase, we encode the pixel with the help of edge responses obtained by the convolution of Robinson compass masks  $(R_m^{(0)}, R_m^{(1)}, R_m^{(2)}, R_m^{(3)}, R_m^{(4)}, R_m^{(5)}, R_m^{(6)}, R_m^{(7)})$  with original face image. For a given pixel, its EALDBP is obtained by encoding the directional and magnitude deviational information of only four edge responses. Since, the Robinson compass mask is symmetric, the magnitudes of edge responses of first four masks  $(R_m^{(0)}, R_m^{(1)}, R_m^{(2)}, R_m^{(3)})$  is similar but with opposite sign with the magnitudes of edge responses of next four masks  $(R_m^{(4)}, R_m^{(5)}, R_m^{(6)}, R_m^{(7)})$ . For instance, the magnitude of edge response of  $R_m^{(2)}$  is similar with the magnitude of edge response of  $R_m^{(6)}$  but with opposite sign. Hence, we have considered only the edge responses of four masks and the remaining are excluded thereby the computational burden also reduces. Moreover, [34] proved that the patterns with high symmetry level occur more frequently in the face images. Due to this property, Robinson compass mask can represent the symmetrical facial features much effectively. We encode the representation by using four directional responses to form a binary code.

The Robinson compass mask is applied over an entire facial image producing a set of response magnitudes corresponding to four directions, as

$$E_R^{(i)} = I \times R_m^{(i)}, \quad 0 \leq i \leq 3 \quad (2)$$

Where  $E_R^{(i)}$  is the edge response and it is obtained after the convolution of  $i^{th}$  Robinson compass mask  $(R_m^{(i)})$  with the original facial image  $I$ . Then we search

for the direction of first top positive response as well as first top negative response. Let's consider a pixel located at a position  $(x, y)$ ,  $d_P(x, y)$  be the positive direction and  $d_N(x, y)$  be the negative direction, they are measured as;

$$d_P(x, y) = \operatorname{argmax}_i \left( E_R^{(i)}(x, y) \right) \quad (3)$$

And

$$d_N(x, y) = \operatorname{argmin}_i \left( E_R^{(i)}(x, y) \right) \quad (4)$$

Here we assign three bits for  $d_P(x, y)$  and another three bits for  $d_N(x, y)$ , hence the total number of bits constructed are 6. For the remaining two bits, we apply LBP process over the edge responses of second top positive and second top negative responses. In this process, initially we search for second top positive and second top negative edge responses and select the corresponding pixels. The magnitude of these two pixels are subtracted from the magnitude of center pixel and based on the obtained values, they are assigned for 0 or 1. In the process of Robinson encoding, for a give center pixel there exists only four neighbor pixels. Among these four pixels, one is  $d_P(x, y)$  and another is  $d_N(x, y)$ . After excluding these two pixels, only two pixels are left. Among these two, we have to find the positive and negative values. Based on the obtained values, the original corresponding pixels are chosen as they are subjected for LBP process, as depicted below;

$$m_P = S(i_P - i_C) \quad (5)$$

And

$$m_N = S(i_N - i_C) \quad (6)$$

Where

$$S(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases} \quad (7)$$

Where  $i_P$  is the corresponding positive pixel in the original facial image,  $i_N$  is the corresponding negative pixel in the original facial image and  $i_C$  is the center pixel in image  $I$ . Then the final EALDBP of a pixel is obtained by the concatenation of all four values as obtained through Eqs. (3)-(6). Let  $EALDBP^R(x, y)$  be the obtained binary code through Robinson Encoding of a pixel located at position  $(x, y)$ , it is obtained as

$$EALDBP^R(x, y) = [d_P(x, y) m_P d_N(x, y) m_N] \quad (8)$$

An example demonstration about the  $EALDBP^R(x, y)$  calculation is shown in the following Fig. 5.

### 3.3.2. Gaussian encoding ( $EALDBP_G^C$ )

Under this encoding, a pixel is represented with an eight bit binary code with the help of edge responses obtained after the convolution of original facial image with eight Gaussian masks. The eight Gaussian masks are derived with the help of  $G_\sigma$  as shown in Eq.(1). Let the eight Gaussian masks are represented as  $(G_m^{(0)}, G_m^{(1)}, G_m^{(2)}, G_m^{(3)}, G_m^{(4)}, G_m^{(5)}, G_m^{(6)}, G_m^{(7)})$ , the reference mask used to derive these masks are  $G_\sigma$ . The Gaussian masks are generated by keeping an angular deviation of  $45^\circ$  between successive masks. The Gaussian compass masks is applied over the entire face image producing a set of edge responses corresponding to the directions, as

$$E_G^{(i)} = I \times G_m^{(i)}, \quad 0 \leq i \leq 7 \quad (9)$$

Where  $E_G^{(i)}$  is the edge response and it is obtained after the convolution of  $i^{th}$  Gaussian compass mask ( $G_m^{(i)}$ ) with the original facial image  $I$ . Then we search for the directions of first top positive response as well as first top negative response. Along with these two directions, we also search for second top positive and second top negative directions to find out the magnitude deviations with center pixel.

Similar to the encoding process employed in Robinson encoding, the first top positive direction is measured with the help of Eq. (3) and the first top negative direction is measured with the help of Eq. (4). Similarly, second top positive and second top negative edge responses and the corresponding pixel's binary patterns are measured through Eq. (5) and (6) respectively. The major difference between the Robinson and Gaussian Encoding is the total number of neighbor pixels present at the calculation of second top positive and negative responses. In the Robinson encoding we have only two pixels while in the Gaussian encoding we have totally six pixels. After finding the second top negative and positive responses, their corresponding pixels are subjected to LBP process. Finally, the pixel is represented with eight bit binary code and this code covers the information of four neighbor pixels, whereas in LDN, only two neighbor pixel's information is conveyed. The final EALDBP of after Gaussian encoding is

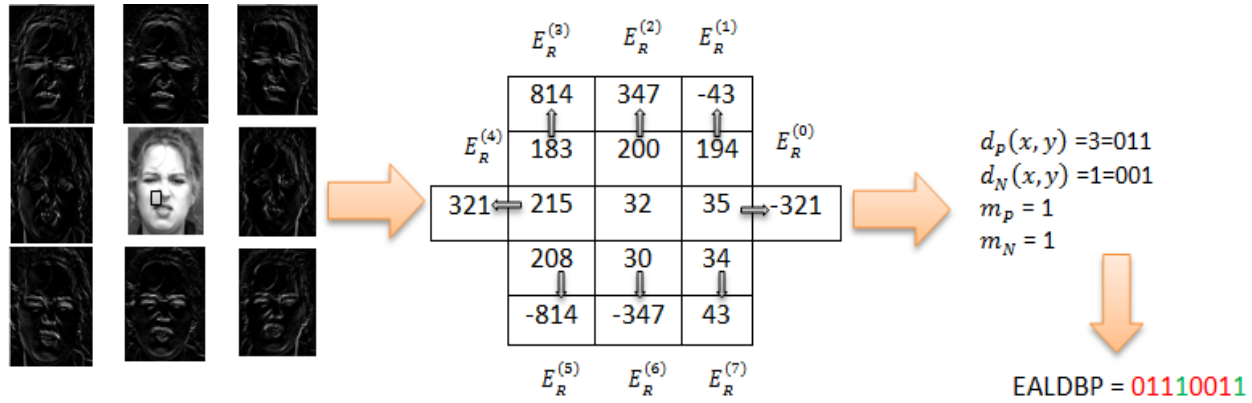


Figure. 5 Robinson encoding

represented as  $EALDBP_{\sigma}^G$  and it is obtained with the help of Eq. (8). In  $EALDBP_{\sigma}^G$ ,  $\sigma$  is the standard deviation of Gaussian mask and in our approach, we have considered  $\sigma = 1$ .

### 3.4 Face descriptor

After the representation of each pixel I the face image with EALDBPs, we applied Histograms to describe the face image. Over the obtained EALDBPs of face image, we measure EALDBP Histogram called as EH. Since the histogram is an efficient descriptor through which we can fetch all the possible features (Ex. edges, textures, corners, and spots), we have used it as a face descriptor. Generally, the Histogram is evaluated over an entire image. Due to this reason, the obtained face descriptor neglects the location information. Hence to aggregate the location information, we divide the face image into small blocks and let it be  $B_1, B_2, \dots, B_N$ . After dividing the encoded image into N small sized blocks, we compute histogram for every block. In this process, we have considered EALDBP encoded face image as input, hence the pixels are represented in the form of binary codes. Thus we use each binary code as histogram bin and then accumulate the total number of such kind of binary codes in the each block. Mathematically the Histogram calculation is represented as;

$$H_i(c) = \sum_{(x,y) \in B_i} \sum_{EALDBP(x,y)=c} a \quad (10)$$

Where  $H_i$  is the histogram of Block  $B_i$ ,  $(x, y)$  is the position of a pixel in the block  $B_i$ ,  $c$  is the EALDBP code,  $EALDBP(x, y)$  is the EALDBP code for the position  $(x, y)$  and 'a' is the accumulation count. Finally, the EH is obtained by the concatenation of Histograms of all blocks. Mathematically the EH is represented as;

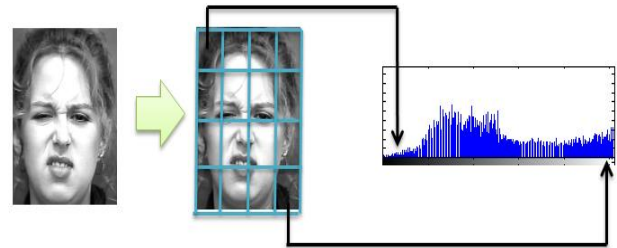


Figure. 6 EH computation

$$EH = \prod_{i=1}^N H_i \quad (11)$$

Where  $\prod$  is the concatenation operator and N is the total number of blocks into which the image is divided. The finally the obtained EH signifies the global face descriptor and the sample demonstration about EH calculation is shown in Fig. 6.

### 3.5 Recognition

We have employed two different machine learning algorithms separately for face expression and expression recognition. They are K-Nearest Neighbor (K-NN) algorithm for face recognition [36] and Support Vector Machine (SVM) [37] Algorithm for expression recognition. Assuming the ideal conditions, the expression recognition is complex compared to the face recognition. Hence we have employed a simple K-NN algorithm for face classification and SVM for expression classification. Moreover, we also described each face with its structural and texture features that provide a sufficient discrimination between faces. In the case of expression recognition, different people may express different expressions in same manner. In such conditions, the projection of data is required to discriminate them effectively. This higher order projection space is possible with SVM; hence we have employed it for expression recognition.



## 4. Experiments

To evaluate the performance of proposed system, we have conducted a series of experiments for both face recognition and face expression recognition. The face recognition performance is evaluated under several varying conditions like pose and illumination. Further, the face expression recognition performance is evaluated under six and seven expressions. The situation experiments are conducted with the help of MATLAB tool.

### 4.1 Face recognition

Under this phase, we tested our proposed method over a standard public dataset; Georgia Tech Face dataset. During the simulation of this dataset, the face region is detected through viola jones algorithm and then every face image is resized to  $200 \times 200$  pixels such that the uniformity is ensured. At the calculation of histogram, every encoded face image is partitioned into several blocks and the size of each block is kept as  $20 \times 20$  and hence the total number of blocks obtained is 100. Next, during the comparison, we compared the results of proposed approach with the results of LBP and LDP.

The Georgia Tech face dataset [38] consists of 750 facial images and they are captured with the help of 50 people in two or three session between 06-01-1999 and 11-15-1999 at the Center for Signal and Image Processing at Georgia Institute of Technology. For every person, there are 15 colored JPEG images with cluttered background and the resolution of each image is kept as  $640 \times 480$  pixels. On an average, the size of only face region of the images in this dataset is  $150 \times 150$  pixels. The images are having frontal and/or tilted faces with different scaling, lighting condition and facial expressions. Every image is labeled manually to determine the face position. Some examples of Georgia Tech face images are show in Fig. 7.

Among the 750 images of Georgia Tech face dataset. We have used 500 mages for Training and remaining 250 images for testing. For every subject, we have considered 10 images thereby the total number of image considered for training are 500. Similarly, for every subject we have considered 5 images for testing such that the total number of testing images is 250. By changing the images in the training, we have conducted a five-fold validation. At every phase of validation, we have compared the proposed EALDBP with several existing methods and the results are shown in Fig. 8.



Figure. 7 Sample images of georgia tech face dataset

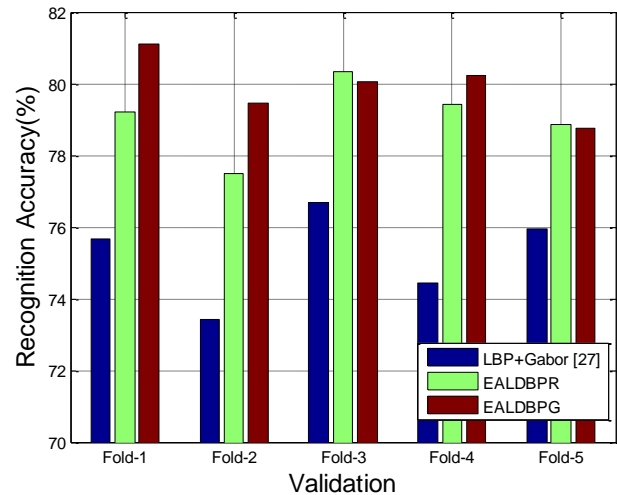


Figure. 8 Recognition accuracy comparison of the EALDBP with existing methods in georgia tech face dataset

Fig. 8 shows the comparison between proposed and existing methods through recognition accuracy. For comparison purpose, we have considered the method proposed in [27], LBP and 2D Gabor wavelet filter for face recognition. In Fig. 8, the maximum recognition accuracy is observed for the proposed EALDBP at all cases of validations. At every validation, the facial images considered for training and testing are random in nature, means the training and testing images are selected purely on a random basis. In this manner, we have simulated for five-fold validation.

At all validations, we can see that the proposed EALDBP has obtained a greater recognition performance because, compared to the Gabor filter, the proposed directional pattern is more effective in the provision of discrimination between facial images. Among the two method of proposed approach such as EALDBPG and EALDBPR, we can see that the maximum recognition accuracy is observed the EALDBP has more accuracy at three fold such as Fold-1, Fold-2 and Fold-4. In the remaining Validations such as FODl-3 and Fold-5, the EALDBPG has obtained good recognition accuracy. From the Fig. 8, we measured the average recognition accuracies; they are 75.2339%, 79.0621% and 79.9152% for LBP + Gabor, EALDBPR and EALDBPG respectively. Further the average recognition accuracy at different validations is

observed as 78.6673%, 76.7809%, 79.0203%, 78.0307% and 77.8580% for Fold-1, Fold-2, Fold-3, Fold-4, and Fold-5 respectively.

#### 4.2 Expression recognition

Under this simulation study, two different facial expression databases are considered. They are (1) (Extended) Cohn-Kanade (CK+) [39] and (2) Japanese Female Face Expression (JAFFE) [40].

CK+ is an extended version of CK database which consist of 486 sequences acquired from 97 subjects. This dataset consists of both posed and non-posed (spontaneous) expressions. Compared to the images present in CK dataset, the sequences of CK+ are increased by 22% and captured with an additional 27% subjects. In this dataset, every sequence begins with a neutral expression and ends with peak expression. The peak expression is fully coded with Facial Action Coding System (FACS) and given an emotion label. The original size of every image is noticed as  $490 \times 640$  and at the simulation every test image is cropped according to the requirements and on an average the size of cropped is noticed as

$310 \times 260$ . All the images are in .PNG format. Some samples of CK+ data set are shown in Fig. 9.

Under the simulation of CK+ dataset, in our setup, we have selected totally 1123 images and each image is labeled with any of the seven basic emotions. To simulate the 7-class expression recognition, the entire 1123 images are categorized into seven expressions (Anger – 169, Contempt – 120, Disgust – 158, Fear – 113, Happy – 213, Sad – 133, and Surprise - 219). Along these expressions we also prepared a neutral set with 855 images. To construct the neural expression set, the first frame from each seuqnece is selected, resulting in a 8-class expression dataset with 855 images. The experimental simulation is done for Five-fold cross validation. For this purpose, we have divided the entire set into five sets. At every validation, among the five partiotions, Four are used to train the classifier and the last one is used for testing. This process is repeated for Five times and at every run the testing group is changed. Here we reported the average recognition results. The

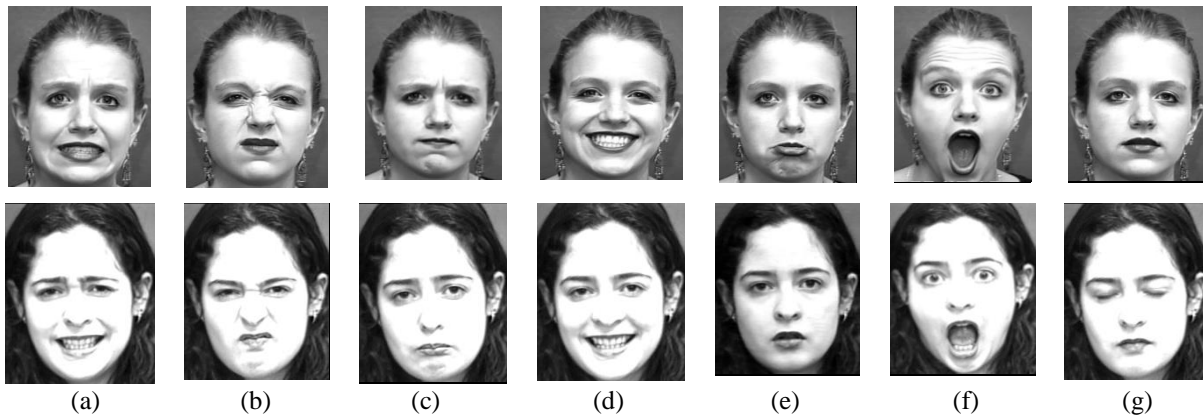


Figure. 9 Samples from CK+ dataset :(a) angry (b) disgust (c) fear (d) happy (e) sad (f) surprise, and (g) neutral

Table. 1 Confusion matrix of 7-class expression recognition in the CK+ dataset

	Angry	Contempt	Disgust	Fear	Happy	Sadness	Surprise	Total
Angry	131	6	7	0	0	25	0	169
Contempt	0	97	0	4	0	12	7	120
Disgust	8	0	142	8	0	0	0	158
Fear	6	0	0	100	0	7	0	113
Happy	0	6	0	8	194	0	3	211
Sadness	9	6	0	5	0	113	0	133
Surprise	0	0	0	2	6	3	208	219
Total	154	115	149	127	200	160	218	1123

classified expression results of every test set are represented through a confusion matrix, as shown in Table 1.

As per the confusion matrix shown in Table 1, the greater performance is observed for Surprise, Happy and Disgust expressions while the lowest performance is observed for Angry and sad expressions. Since the surprise, happy and disgust expressions have unique facial variations, almost all the input test images are recognized correctly. For Surprise expression, the opened mouth is a unique identification mark and hence it has gained higher TPR when compared to remaining expressions. Next, the happy expression and Disgust also have unique identification marks; they are Rise in Cheeks and Skin folding between the two eyebrows respectively. The proposed approach mainly focused on the discovery of edge directions in which the movements are incurred thereby providing a perfect discrimination between all the expressions. The EALDBPs obtained for Happy, Disgust and Surprise are too different and hence they have recognized more accurately. Next, worst confusion is occurred among the sadness, contempt and anger expressions. The main reason is that they have similar characteristics and it is somewhat tough to recognize

them through a single image/frame. Further From Fig.11, we can see that the Recall of Surprise expression is 94.9897% and for Angry expression it is of 77.5120%. Next the maximum Precision is obtained for Happy expression and it is observed as 97.3321% while the minimum Precision is gained for sadness expression and it is approximately 70.6312%. Next, the JAFFE dataset is a facial expression dataset and it is captured through the facial expressions of only Japanese female subjects. JAFFE contains 213 images of 7 different facial expressions such as Neutral, Angry, Disgust, Fear, Happy, Sad and Surprise. All these images are captured with the help of 10 Japanese models. In every image, the face is posed in front al view and the candidate's hair is tied back to represent all the expressive landmarks of her face. The head in each image is usually in frontal pose, and the subject's hair was tied back to expose all the expressive zones of her face. The original size of each image in this dataset is noticed as  $256 \times 256$  and at the simulation every test image is cropped and on an average the size of cropped is noticed as  $170 \times 200$ . All the images are in .TIFF format. Some samples of JAFFE data set are shown in Fig. 10.

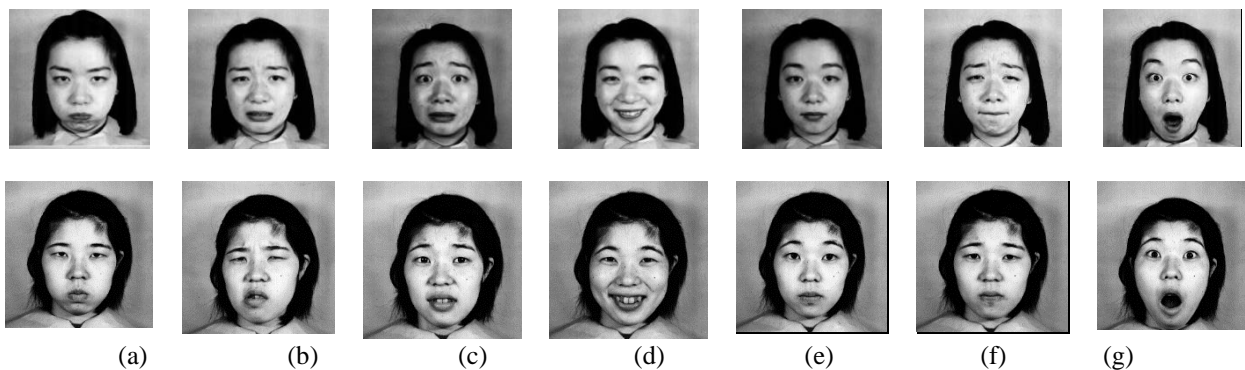


Figure. 10 Samples from JAFFE dataset (a) angry (b) disgust (c) fear (d) happy (e) neutral (f) sad, and (g) surprise

Table. 2 Confusion matrix of 7-class expression recognition in the JAFFE dataset

	Angry	Disgust	Fear	Happy	Neutral	Sadness	Surprise	Total
Angry	23	0	3	0	0	4	0	30
Disgust	0	27	2	0	0	0	0	29
Fear	0	5	25	0	0	2	0	32
Happy	0	0	0	29	0	0	2	31
Neutral	2	0	0	0	26	2	0	30
Sadness	4	0	4	0	0	23	0	31
Surprise	0	0	0	0	0	0	30	30
Total	29	32	34	29	26	31	32	213

In the simulation of JAFFE dataset, we have considered seven expressions. The validation is done in three fold. The images considered for simulation are initially divided into three sets. This is done based on availability of multiple images of same expression and same person. On an average, each subject has three images of same expression under every category. Out of these three images, at every phase of validation, two images are trained and one image is tested. This process is repeated for multiple times by changing the testing images and the overall recognized results are shown in Table 2. From this table, we can see that the maximum performance is achieved for Surprise expression while the minimum performance is achieved for Sadness expression. In this simulation, we have added an additional neutral expression and hence some of the other faces (Sad) are recognized as neutral due to the similar characteristics between them. Next, some of the disgust and fear expression also have similar facial characteristics hence they are recognized as opposite expressions. Some of the happy face expressions are recognized as Surprise due to the similar mouth movements. Based on the values shown in confusion matrix, the maximum Recall is observed as 100% for surprise expression and maximum precision is observed as 100% for Happy and Neutral expressions. Further from Fig. 11, we can see that the minimum recall is observed as 74.1920% for sadness expression and minimum Precision is observed as 73.5312% for Fear expression.

### 4.3 Comparison

To alleviate the robustness of proposed approach, we have compared it with several existing methods such as LBP [11], Gabor with LBP [27], LMBP [22], and LDN with DGLTP [28]. The comparison is done only with the methods those have used CK and JAFFE for simulation purpose. Further we considered only the methods those focused mainly on the encoding pattern, means they have discovered a new coding strategy. The method proposed in [11] considered LBP for face representation and SVM for classification. The LBP process encodes a pixel by considering only the magnitude deviations between center pixel and its neighbor pixels. However, the main drawback is loss of information.

The method proposed in [27] applied LBP over the sub bands obtained through Gabor wavelet transform. Even through Gabor wavelet transform is affecting in ensuring pose invariance, the sub bands discard the hue information loss and also break the correlation between pixels. Compared to this method, the proposed method is effective because it has

applied directional coding by which the movements of facial expression are effectively captured.

Next, in the LMBP [22], the encoding process is carried out based on the mean of a local neighborhood, means the pixel encoding is done by comparing the neighbor pixels with the mean of a local neighborhood. However, this method is not effective in the case of sudden changes in the pixel intensities because for a local neighborhood which contains smooth as well edges, the mean value is too abnormal and it creates a more confusion between smooth region pixels and edge region pixels. This problem is solved in the proposed approach due to the consideration of directional information.

Next, focusing over the directionality, I. Michael Revina, W.R. Sam Emmanuel [28] proposed LDN followed by DGLP to extract the directionality information. For example, in the contempt emotion, only right or left side of lips (in the closed format) is moved up, means the gradient is in the upward direction. To recognize such types of emotions, directionality information also needs to be included in the LBP. LDN discovers the directionality and integrates it into the feature vector. However, the directionality discovery at single orientation is not efficient to recognize emotion in multiple orientations. Moreover, the LDN encoded the pixel through only directional information means for a given center pixel, only two of its neighbor pixels are included and remaining pixels are excluded, results in a heavy information loss. Hence the proposed approach considered directional and magnitude information thereby it considered four neighbor pixels and reduced the information loss. Hence the proposed approach has gained higher recognition performance compared to all the existing methods.

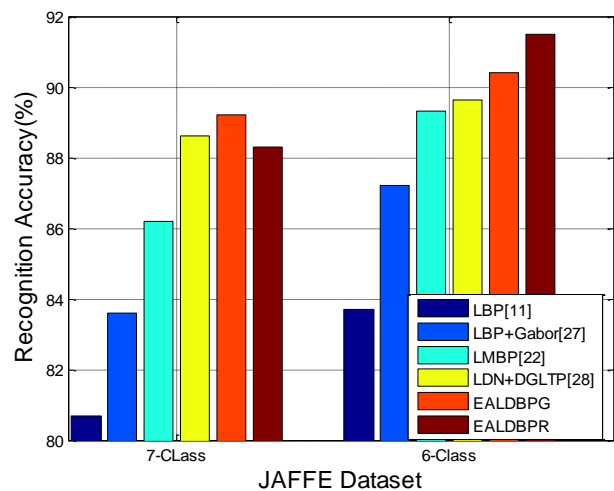


Figure. 12 Accuracy comparison between proposed and existing methods in the JAFFE dataset

Fig. 12 provides a comparison between proposed and earlier specified existing methods through the accuracy obtained over the simulation of JAFFE dataset. In the simulation, two different instances of simulations are evaluated they are 7-class expression and 6-class expression. In the former one, totally seven expression (Angry, Disgust, Fear, Happy, Neutral, Sad, and Surprise) are considered while in the second one only six expressions (Angry, Disgust, Fear, Happy, Sad, and Surprise) are considered. Next, under the simulation through proposed model, we have considered individual masks and verified the performance. Based on the obtained results, we can see that the EALDBP with Gaussian compass mask (ELADBPG) has gained higher accuracy (89.2254%) in 7-class and EALDBP with Robinson mask (ELADBPR) has gained higher accuracy (91.5542%) in 6-class.

The average recognition accuracy of proposed approach under JAFFE dataset is observed as 82.2000%, 85.4000%, 87.5000%, 89.1300%, 89.8000% and 89.9000% for LBP, LBP + Gabor, LMBP, LDN + DGLTP, EALDBPG and EALDBPR respectively. From these values we can states that the proposed approach has gained an efficient recognition accuracy compared to the existing methods.

Fig. 13 provides a comparison between proposed and several existing methods through the accuracy obtained over the simulation of CK+ dataset. Here also we have considered the 7-class and 6-class to verify the performance of proposed approach.

In the 7-class expression, we have considered Angry, Contempt, Disgust, Fear, Happy, Sad, and Surprise while in the six-class expression, we have excluded contempt. Based on the obtained results, we can see that the ELADBPG has gained higher

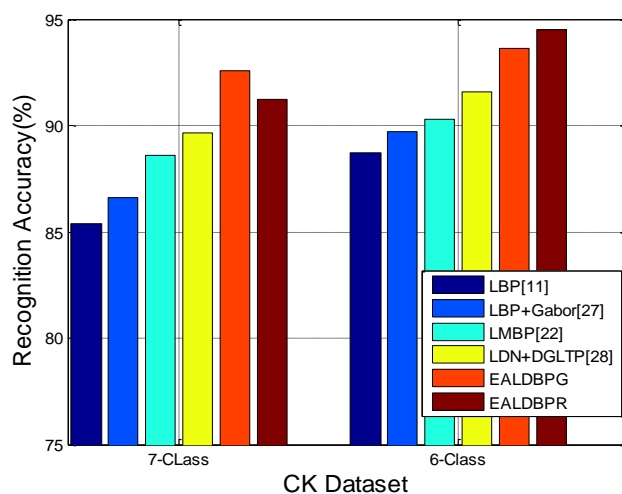


Figure. 13 Accuracy comparison between proposed and existing methods in the CK dataset

accuracy (92.6321%) in 7-class and ELADBPR has gained higher accuracy (94.5523%) in 6-class. The average recognition accuracy of proposed approach under CK+ dataset is observed as 87.5021%, 88.1542%, 89.4563%, 90.6152%, 93.1202% and 92.8542% for LBP, LBP + Gabor, LMBP, LDN + DGLTP, EALDBPG and EALDBPR respectively. From these values we can states that the proposed approach has gained an efficient recognition accuracy compared to the existing methods.

Further, we noticed that the recognition accuracy in CK+ dataset is relatively higher than the JAFFE dataset. The main reason behind this less accuracy is that some of the expressions in JAFFE dataset are much similar with other expressions.

## 5. Conclusion and future scope

In this paper, we have developed a new face encoding scheme called as EALDBP that encodes the facial texture structure into a compact binary code. EALDBP uses directional as well as magnitude information that is more stable to noise and illumination variations. Further, EALDBP employed to different compass masks; they are Robinson compass mask and Gaussian Compass mask to measure the edge responses thereby the directional information is extracted. The obtained compact binary code is an eight bit binary code that is composed two directional pixel's information and two magnitude pixel's information. EALDBP is more effective than LBP and LDP because it conveys more information about a facial expression. Experimental simulations are carried out two datasets such as JAFFE and CK and the obtained results are compared with several existing methods. From the results we have observed that the Gaussian mask is more robust for noise and illumination variations and the Robinson mask is robust to identify the features of expression's structure. Further to found the extra class effect, we have simulated the proposed model in two different scenarios; they are 7-class expression and 6-class expressions. In both cases, the proposed method had shown an outstanding performance than existing methods. The average recognition accuracy of proposed method for CK+ dataset is observed as 93.5922% while for JAFFE dataset it is observed as 90.3898%.

Micro expressions are special type of facial expressions which has very less span of time occurrence and it is approximately 0.5 second. Further they are rapid in nature and the intensity of these short-lived micro-expressions is very low in terms of facial muscles' movement. Hence we have

intended to work over micro expression recognition as our future work.

## References

- [1] R. Cowie, E. D. Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotional expression recognition in human-computer interaction", *IEEE Signal Processing Magazine*, Vol.18, No. 1, pp. 32-80, 2001.
- [2] J. Kumaria, R. Rajesha, and K. M. Poojaa, "Facial expression recognition: A survey", In: *Proc. of Second International Symposium on Computer Vision and the Internet*, Aluva, Kochi, Kerala, India, pp.486 – 491, 2015.
- [3] Erin B. McClure, K. Pope, A. J. Hoberman, D. S. Pine, E. Leibenluft, "Facial expression recognition in adolescents with mood and anxiety disorders", *American Journal of Psychiatry*, Vol. 160, No. 6, pp. 1172–1174, 2003.
- [4] S. Wallace, M. Coleman, A. Bailey, "An investigation of basic facial expression recognition in autism spectrum disorders", *Cognition and Emotion*, Vol. 22, No. 7, pp. 1353–1380, 2008.
- [5] Z. Zeng, M. Pantic, G. I. Roisman, T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 31, No. 1, pp. 39–58, 2009.
- [6] N. Wang, X. Gao, D. Tao, H. Yang, X. Li, "Facial feature point detection: A comprehensive survey", *Neurocomputing*, Vol. 275, pp. 50–65, 2018.
- [7] B. Fasel and J. Luetten, "Automatic facial expression analysis: a survey", *Pattern Recognition*, Vol. 36, No. 1, pp. 259-275, 2003.
- [8] N. G. Bourbakis and P. Kakumanu, "Skin-based face detection-extraction and recognition of facial expressions", *Applied Pattern Recognition*, Vol. 91, pp. 3–27, 2008.
- [9] A. Cheddad, D. Mohamad, and A. A. Manaf, "Exploiting voronoi diagram properties in face segmentation and feature extraction", *Pattern Recognition*, Vol. 41, No. 12, pp. 3842–3859, 2008.
- [10] X. Xie and K. M. Lam, "Facial expression recognition based on shape and texture", *Pattern Recognition*, Vol. 42, No. 5, pp. 1003–1011, 2009.
- [11] B. Singh, R. K. Sharma, R. K. Saxena, and Ms. Ragini Malviya, "Facial Expressions Recognition Based Using LBP and SVM Classifier", *International Journal of Innovative Research in Science, Engineering and Technology*, Vol. 8, No. 8, pp. 8508-8517, 2019.
- [12] M. B. S Divya and N. B Prajwala, "Facial Expression Recognition by Calculating Euclidian Distance for Eigen Faces Using PCA", In: *Proc. of International Conf. on Communication and Signal Processing (ICCSP)*, Chennai, India, pp. 1-4, 2018.
- [13] Zhan Wang, Qiuqi Ruan, and Gaoyun A, "Facial Expression recognition using sparse local fisher discriminant analysis", *Neurocomputing*, Vol. 174, Part B, pp. 756-766, 2016.
- [14] S. Jaiswal, Dr. (Smt.) S. S. Bhadauria, R. S. Jadon, "Comparison Between Face Recognition Algorithm-Eigenfaces, Fisher faces And Elastic Bunch Graph Matching", *Journal of Global Research in Computer Science*, Vol. 2, No. 7, pp.187-193, 2011.
- [15] J. K. J. Julina, T. S. e Sharmila, "Facial Emotion Recognition in Videos using HOG and LBP", In: *Proc. of 4th International Conf. on Recent Trends on Electronics, Information, Communication & Technology (RTEICT)*, Bangalore, India, pp. 56-60, 2019.
- [16] N. T. Cao, A. H. Ton hat, and H. I Choi, "Facial Expression Recognition Based on Local Binary Pattern Features and Support Vector Machine", *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 28, No. 6, pp. 1456012, 2014.
- [17] N. Kauser and J. Sharma, "Facial expression recognition using LBP template of facial parts and multilayer neural network", In: *Proc. of International Conf. on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Palladam, India, pp. 445-449, 2017.
- [18] T. Jabid, M. H. Kabir, and O. Chae, "Robust facial expression recognition based on local directional pattern", *ETRI Journal*, Vol. 32, No. 5, pp. 784–794, 2010.
- [19] T. Xu, J. Zhou, and Y. Wang, "A Variation of Local Directional pattern and its applications for Facial Expression Recognition", In: *Proc. of International Conf. on Signal Processing, Image Processing, and Pattern Recognition*, Jeju Island, Korea, pp. 36-47, 2011.
- [20] M. H. Kabir, T. Jabid, O. Chae, "Local directional pattern variance (LDPV): a robust feature descriptor for facial expression recognition", *The International Arab Journal of Information Technology*, Vol. 9, No. 4, pp. 382-391, 2012.
- [21] P. Carcagn, M. D. Coco, M. Leo, and C. Distanto, "Facial expression recognition and histograms

- of oriented gradients: A comprehensive study”, *SpringerPlus*, Vol. 4, No.1, pp.1-25, 2015.
- [22] M. M. Goyani and N. Patel, “Recognition of facial expressions using local mean binary pattern”, *Electronic Letters on Computer Vision and Image Analysis*, Vol. 16, No. 1, pp. 54–67, 2017.
- [23] B. Zhang, Y. Gao, S. Zhao, and J. Liu, “Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor”, *IEEE Transactions on Image Processing*, Vol. 19, No. 2, pp. 533–544, 2010.
- [24] A. Ramirez Rivera, R. Castillo, and O. Chae, “Local directional number pattern for face analysis: Face and expression recognition”, *IEEE Transactions on Image Processing*, Vol. 22, No. 5, pp. 1740–1752, 2013.
- [25] M. Guo, X. Hou, Y. Ma, “Facial expression recognition using ELBP based on covariance matrix transform in KLT”, *Multimedia Tools and Applications*, Vol. 76, pp. 2995–3010, 2017.
- [26] M. Abdul and R. S. Holambe, “Local binary patterns based on directional wavelet transform for expression and pose invariant face recognition”, *Applied Computations and Informatics*, Vol. 15, No. 2, pp. 163-171, 2017.
- [27] Qian Zhang, Haigang Li, Ming Li and Lei Ding, “Feature extraction of face image based on LBP and 2-D Gabor wavelet transform”, *Mathematical Biosciences and Engineering*, Vol. 17, No. 2, pp. 1578-1592, 2019.
- [28] I. M. Revina, and W. R. Sam Emmanuel, “Face expression recognition using LDN and Dominant Gradient Local Ternary Pattern descriptors”, *Journal of King Saud University – Computer and Information Sciences*, Vol. 6, pp. 1-7, 2018.
- [29] Q. Y. Zhao, B. C. Pan, J. J. Pan, and Y. Y. Tang, “Facial Expression Recognition Based on Fusion of Gabor and LBP Features”, In: *Proc. of International Conf. on Wavelet Analysis and Pattern Recognition*, Hong Kong, China, pp. 52-59, 2008.
- [30] X. Tan and B. Triggs, “Enhanced local texture feature sets for face recognition under difficult lighting conditions”, *IEEE Transactions on Image Processing*, Vol. 19, No. 6, pp. 1635–1650, 2010.
- [31] Viola, P., and Jones, M., “Robust real-time face detection”, *International Journal of Computer Vision*, Vol. 57, No.2, pp. 137–154, 2004.
- [32] G. S. Robinson, “Edge detection by compass gradient masks”, *Computer Graphics and Image Processing*, Vol. 6, No. 5, pp. 492–501, 1977.
- [33] Getreuer, and Pascal, “A Survey of Gaussian Convolution Algorithms”, *Image Processing on Line*, Vol. 3, pp. 286–310, 2013.
- [34] O. Lahdenoja, M. Laiho, and A. Paasio, “Reducing the feature vector length in local binary pattern based face recognition”, In: *Proc. of IEEE International Conf. on Image Processing, (ICIP 2005)*, Genova, Italy, pp. II–914, 2005.
- [35] T. Zhang, Y. Y. Tang, B. Fang, Z. Shang, and X. Liu, “Face recognition under varying illumination using gradient faces”, *IEEE Transactions on Image Processing*, Vol. 18, No. 11, pp. 2599–2606, 2009.
- [36] X. Wang, Z. Chen, and Z. Lin, “Class-Nearest Neighbor Classifier for Face Recognition”, In: *Proc. of International Conf. on Computer Sciences and Applications*, Wuhan China, pp. 325-328, 2013.
- [37] C. W. Hsu and C. J. Lin, “A comparison of methods for multiclass support vector machines”, *IEEE Transactions on Neural Networks*, Vol. 13, No. 2, pp. 415–425, 2002.
- [38] L. Chen, H. Man and A. V. Nefian, “Face recognition based multi-class mapping of Fisher scores”, *Pattern Recognition*, Vol. 38, No. 6, pp. 799-811, 2005.
- [39] P. Lucey, J. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, “The extended Cohn–Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression”, In: *Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. Workshops*, San Francisco, CA, USA, pp. 94–101, 2010.
- [40] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, “Coding facial expressions with Gabor wavelets,” In: *Proc. of 3rd IEEE Int. Conf. Automatic Face Gesture Recognition.*, Nara, Japan, pp. 200–205, 1998.
- [41] K. A. El Dahshan, E. K. Elsayed, A. Aboshoha, and E. A. Ebeid, “Recognition of Facial Emotions Relying on Deep Belief Networks and Quantum Particle Swarm Optimization”, *International Journal of Intelligent Engineering and Systems*, Vol. 13, No. 4, pp. 90-101, 2020.