



# A novel approach for facial expression recognition using local binary pattern with adaptive window

Durga Ganga Rao Kola<sup>1</sup>  · Srinivas Kumar Samayamantula<sup>1</sup>

Received: 13 November 2019 / Revised: 4 July 2020 / Accepted: 18 August 2020 /

Published online: 12 September 2020

© Springer Science+Business Media, LLC, part of Springer Nature 2020

## Abstract

Facial Expression Recognition (FER) is an important area in human computer interaction. FER has different applications such as analysis of student behaviour in virtual class room, driver mood detection, security systems, and medicine. The analysis of facial expressions is an interesting and exciting problem. Feature extraction plays important role in any FER system. Local Binary Pattern (LBP) and its variants are popular for feature extraction due to simplicity in computation and monotonic illumination invariant property. However, the performance of LBP is poor in the presence of noise. This work proposes a novel approach for feature extraction to improve the performance of the FER. In this approach, the LBP is calculated considering 4-neighbors and diagonal neighbours separately. Further, for affective feature description, the concept of adaptive window and averaging in radial directions is introduced. This approach reduces the length of the feature vector as well as immune to noise. Support Vector Machine (SVM) is considered for classification. Recognition rate and confusion matrix are used to assess the performance of the proposed algorithm. Extensive experimental results on JAFFE, CK, FERG and FEI face databases show significant improvement in recognition rate compared to the available techniques both in noise free and noisy conditions.

**Keywords** Adaptive window · Facial expression recognition · Feature extraction · Local binary pattern · Support vector machine

## 1 Introduction

Facial expressions of human beings are universal across all cultures and these are immediate means to understand the mood of the person [23]. It is observed that 55% of communication is easily conveyed to human beings by facial expressions only [14]. Over the past

---

✉ Durga Ganga Rao Kola  
ganga\_408@yahoo.co.in

Srinivas Kumar Samayamantula  
samayamantula1963@gmail.com

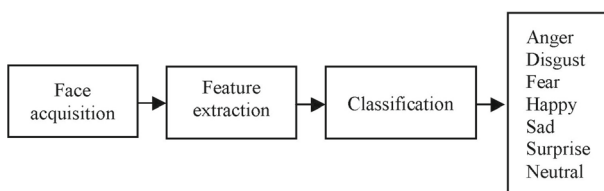
<sup>1</sup> Department of ECE, UCEK, JNTUK, Kakinada, 533003, India

decade, Facial Expression Recognition (FER) has been a significant area in the field of computer vision [57]. The FER has many exciting applications in human-computer interaction, driver mood detection, and data driven animation [43]. With the development of technology, the analysis of facial expressions may be applied to education field, where we can able to understand the student commitment in case of virtual class room. The FER can be applied to medicine in which patient pain has been detected automatically through online by the doctor [2]. Analysis of facial expressions can also be applied for security to detect suspicious persons by observing their emotions. Even in animation, the FER is used to detect the emotions from given image [44]. Hence, the research area of FER is very interesting and exciting in many practical applications [46].

The basic FER system is shown in Fig. 1. It mainly consists of three fundamental steps: face acquisition, feature extraction, and classification. Face acquisition involves the detection of the face region from the given image or video sequence. After detecting the face region, selected face area is cropped to eliminate the background part in the image. The cropped face area is further normalized to a fixed size. Once the face region is selected, then features are extracted from this region. The performance of any FER system depends on an efficient extraction of significant features from the face [47]. Feature extraction is helpful for interpreting meaningful information from facial image [20]. A detailed survey on fundamentals of image features and descriptors found in [4]. Any extracted feature should satisfy three important properties: first, it must be high discriminative among the inter class variations and low discriminative within the same class variations. Second, the procedure of feature computation should be simple and it should be described in a low dimensional space. Third, the extracted feature must be more robust against illumination variation and noise.

The final step of FER system is to choose the classifier for features extracted in the previous step. The features are fed into the selected classifier which performs classification based on the similarities between the extracted features. The widely accepted basic emotions are: neutral, anger, disgust, fear, happy, sad, and surprise. Various classifiers include K-Nearest Neighbours (KNN), Neural Networks (NN) and Support Vector Machines (SVM). It has been proven that for FER, the SVM performs better than other classifiers.

In this paper, a new approach for feature extraction by applying Local Binary Pattern (LBP) separately to 4-neighbors and diagonal neighbors is considered instead of applying on eight neighbours of a central pixel like traditional LBP. This approach reduces the feature length from 256 to 32, thereby the computational complexity is reduced. In order to have noise robustness, a new neighbourhood is defined by averaging along radial direction for texture classification [42]. In this work, an adaptive window is considered around each pixel based on the intensity variations. Further, in order to achieve the robustness in presence of noise, averaging of pixel intensities in each radial directions is considered, thereby a  $3 \times 3$  window is used to calculate the LBP.



**Fig. 1** Facial expression recognition system

The remainder of the paper is as follows: Related work of FER is described in Section 2, a brief review on LBP and its variants is discussed in Section 3. The proposed method for feature extraction is given in Section 4. Experimental setup is presented in Section 5. Simulation results are discussed in Section 6. Finally, concluding remarks are given in Section 7.

## 2 Related work

Feature extraction plays an important role in any FER system [5]. The extracted features from the face must have high discrimination among the different classes but less discrimination within the same classes. There are two types of approaches for feature extraction: geometric based methods and appearance based methods [6]. In the geometric based approach, the feature vector is formed by extracting the locations of the facial components and shape. However, in real time, the extraction of geometric features is difficult due to large changes in faces across the different cultures [10]. So, appearance based approach gained popularity in the feature extraction [48]. In appearance based methods, spatial filters are applied on either the whole face or specific region of the face to obtain the texture information [13]. Gabor wavelets and LBP are two popular feature extraction techniques in appearance based methods [8]. Gabor wavelets extract the features in multiple scales and orientations but computational complexity is high.

The LBP was a powerful texture descriptor for feature extraction in appearance based methods [38]. The LBP was used in various areas like content based image retrieval, texture analysis, face recognition, bearing fault diagnosis and many more [30]. LBP based methods were computationally simple and efficient, but they were affected by the noise. In order to improve the robustness of LBP, different variants of LBP were proposed for feature extraction. Tan and Triggs [51] proposed Local Ternary Pattern (LTP), which partially solved the problem in the presence of noise but its threshold selection was difficult. Jabid et al. [25] proposed Local Directional Pattern (LDP) which was based on the edge responses in eight different directions and then binary code was formed based on these responses. The performance of LDP depends on the number of edge response directions [26]. Tong et al. [53] proposed Local Gradient Coding using Horizontal and Diagonal principle (LGC-HD) operator to reduce the feature length, but it did not consider the effect of noise. Holder and Tapamo [21] proposed Improved Gradient Local Ternary Pattern (IGLTP) using Scharr operator to reduce the effect of variation in illumination, but the selection of threshold was difficult. Yee et al. [55] proposed completed Local Ternary Pattern (CLTP) to overcome the drawbacks in basic LBP operator. Fan et al. [18] proposed Structure Co-Occurrence Texture (Scoot) that considers both spatial structure and co-occurrence texture statistics. Histogram of Oriented Gradients (HOG) was another descriptor used for feature extraction in FER system. Eng et al. [16] applied HOG for feature extraction in FER system. Nigam et al. [40] combined the HOG and wavelet transform for feature extraction to improve the performance of FER system. All these methods have partially solved the effect of the noise on feature extraction.

Recently, Convolutional Neural Networks (CNN) and deep learning successfully applied in the applications of computer vision. The main advantage of deep learning was that it integrates both feature extraction and classification. Cho et al. [11] proposed Relational Graph Module (RGM) for heterogeneous face recognition. It mainly describes the high-level relational information between facial components. Bi et al. [7] proposed conditional

Generative Adversarial Network (cGAN) for Face Sketch Synthesis. cGAN was used to learn the mapping relationships between the face and responding sketch. Huang et al. [24] proposed an adaptive curriculum learning loss for deep face recognition, which automatically emphasizes easy samples first and hard samples later. Pitaloka et al. [45] proposed Enhanced CNN for automatic emotion recognition. They showed that pre-processing methods like resizing, face detection, cropping etc., influence the performance of CNN. Chen et al. [9] proposed deep CNN based on edge computing. It overcomes the problem of high similarity and imbalanced number of sample facial expressions. Kim et al. [33] proposed Deep Neural Network (DNN) structure by combining appearance and geometric features. Li et al. [35] presented faster regions with CNN features to overcome the problems of low level data operation and complexity in feature extraction. Yang et al. [54] proposed Weighted Mixture Deep Neural Network (WMDNN) for effective feature extraction. They used gray scale images and their corresponding LBP images to improve the performance of the FER system. Shan et al. [50] proposed automatic FER using deep CNN network structure. Jung et al. [28] developed the FER system using both CNN and DNN and demonstrated experimentally that CNN outperform compared to DNN. Liliana [36] proposed Deep CNN by detecting the occurrence of facial Action Units for efficient FER. Aneja et al. [3] proposed DeepExpr model for FER by using CNN and transfer learning technique for stylized animated characters. Minaee and Abdolrashidi [39] proposed deep learning approach based on attentional convolutional network which was able to focus on the important parts of the face. Jaiswal and Nandi [27] developed CNN architecture for real-time emotion detection system.

Very recently, CNNs employed for Salient Object Detection (SOD). The SOD helps in determining the most visually characteristic objects in an image. Fan et al. [17] evaluated the recent CNNs based SOD models and further they proposed high quality Salient Objects in Clutter (SOC) dataset. Zhang et al. [56] proposed uncertainty network named UC-Net for RGB-D inspired by human uncertainty in ground truth for saliency detection. Zhao et al. [59] proposed an Edge Guidance network (EGNet) for SOD. EGNet models complementary features: salient object and edge information. Zhang et al. [58] presented a new weakly-supervised SOD method and introduced a new scribble based saliency dataset S-DUTS. Deep learning was applied for many applications in computer vision. However, some limitations of deep learning are: it requires large number of training samples, convergence to local minimum and huge computation compared to conventional feature extraction techniques. Hence, it is proposed a novel approach for FER by using LBP with the concept of adaptive window.

### 3 LBP and its variants

In this section, the basic theory of the LBP and its variants such as LTP, LGC-HD operator, and IGLTP are discussed.

#### 3.1 LBP

The LBP is an efficient texture descriptor and it is used in many areas like image retrieval, face recognition, and remote sensing, etc [22]. The main advantage of LBP lies with its less computational complexity and illumination invariant property [30]. The LBP operator for a

given window having  $P$  number of pixel intensities and radius  $R$ , is calculated as follows.

$$LBP_P^R(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \tag{1}$$

$$s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases} \tag{2}$$

where,  $g_c$  is the intensity value of the center pixel  $(x_c, y_c)$  and  $g_p$  denotes the gray values of neighbours around centre pixel with radius  $R$ . For instance, the window shown in Fig. 2 has  $P = 8$  and  $R = 1$ , and the LBP produces 256 different output values [41]. The LBP operator is applied on each pixel to compute the LBP codes [1]. The 256-bin histogram is formed from LBP codes and it is used as a feature vector for FER [49]. The main drawback of LBP is that its code changes if central pixel value changes due to noise or any other disturbance.

### 3.2 LTP

The LTP was proposed by Tan and Triggs [51] to improve the noise sensitivity of the central pixel in LBP. The LBP is extended into three valued codes: -1, 0, and 1 instead of 0 and 1. The LTP can be calculated by replacing  $s(x)$  in (2) by

$$s(g_p, g_c, t) = \begin{cases} 1, & g_p \geq g_c + t \\ 0, & |g_p - g_c| < t \\ -1, & g_p \leq g_c - t \end{cases} \tag{3}$$

where,  $t$  is a threshold. The LTP uses binary codes separately for positive and negative values because of large range with three values. The histogram is calculated for both codes and they are combined to get a final feature. The LTP partially solved the problem of noise sensitivity in LBP, but the selection of threshold was difficult and further, the feature length was more when compared to the LBP.

### 3.3 LGC-HD operator

The LGC-HD operator was proposed by Tong et al [53]. This operator was based on local gradient coding in which horizontal gradients and diagonal gradients were considered. This operator has an advantage of less computational complexity, but powerful discriminative operator for texture changes in face. For a given window shown in Fig. 2, the LGC-HD operator is calculated as follows:

$$LGC-HD = s(g_1 - g_3)2^4 + s(g_4 - g_5)2^3 + s(g_6 - g_8)2^2 + s(g_1 - g_8)2^1 + s(g_3 - g_6)2^0 \tag{4}$$

Fig. 2 The 3 × 3 window

$g_1$	$g_2$	$g_3$
$g_4$	$g_c$	$g_5$
$g_6$	$g_7$	$g_8$

where,  $s(x)$  is defined in (2). The LGC-HD operator is applied on each and every pixel of a given image and LGC-HD codes are calculated. The histogram of all LGC-HD codes is used as a feature vector.

### 3.4 IGLTP

The IGLTP was proposed by Holder and Tapamo [21]. This operator was based on the Scharr operator and the LTP. First, the Scharr operator was used to calculate the gradient magnitude of a given image. The kernels used to compute gradient by using Scharr operator for a given  $3 \times 3$  window are shown in Fig. 3. These are applied on image to compute the gradients in horizontal ( $G_x$ ) and vertical directions ( $G_y$ ). Finally, the gradient magnitude is calculated by using the formula:

$$G_{x,y} = \sqrt{G_x^2 + G_y^2} \tag{5}$$

After calculating the gradient magnitude, the LTP is applied on this gradient image to get the feature vector.

## 4 Proposed method

In this section, the modified LBP for feature extraction method using adaptive window approach to improve the noise robustness is presented.

### 4.1 Feature extraction using proposed method

Feature extraction plays a vital role in the performance of FER system. Two important characteristics should be considered in feature extraction, viz., size of the feature vector and robustness in the presence of noise [32]. This has motivated to propose a new approach to calculate the LBP to have both these advantages. In this approach, LBP is calculated considering 4-neighbors and diagonal neighbours separately. Further, the size of the window is adaptively chosen based on the variance of the window. If the size of the window is more than  $3 \times 3$ , then the averaging of pixel intensities in the same radial direction is considered to reduce the size of the effective window to  $3 \times 3$ . This increases the noise immunity and also becomes invariant to illumination variations. Consider a  $3 \times 3$  window as shown in Fig. 2. The steps of the proposed method for feature extraction are as follows:

$$\begin{matrix} \begin{bmatrix} -3 & 0 & +3 \\ -10 & 0 & +10 \\ -3 & 0 & +3 \end{bmatrix} & \begin{bmatrix} -3 & -10 & -3 \\ 0 & 0 & 0 \\ -3 & +10 & +3 \end{bmatrix} \\ \text{(a)} & \text{(b)} \end{matrix}$$

Fig. 3 Scharr kernels for  $3 \times 3$  mask: **a** horizontal **b** vertical

- (i) Apply the LBP by considering only horizontal and vertical neighbors and calculate the histogram over the entire image by using (6) and this feature vector is represented as  $f_1$ .

$$LBP-HV = s(g_2 - g_c)2^3 + s(g_5 - g_c)2^2 + s(g_7 - g_c)2^1 + s(g_4 - g_c)2^0 \quad (6)$$

- (ii) Apply the LBP by considering only diagonal neighbors and calculate the histogram over the entire image by using (7) and this feature vector is represented as  $f_2$ .

$$LBP-D = s(g_1 - g_c)2^3 + s(g_3 - g_c)2^2 + s(g_8 - g_c)2^1 + s(g_6 - g_c)2^0 \quad (7)$$

where,  $s(x)$  is defined in (2).

- (iii) Now, a new feature vector is defined by concatenating the above two feature vectors as follows:

$$f = [f_1, f_2] \quad (8)$$

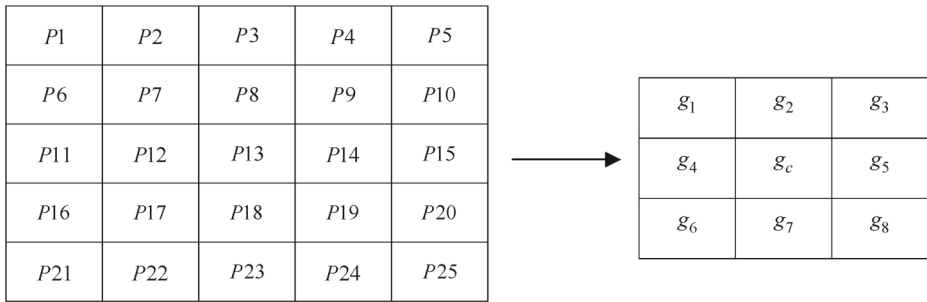
In this approach, each operator considers only four neighbours at a time, hence the feature length of each operator is 16 and the concatenated feature has the length of 32 only, which is less in length compared to the traditional LBP. However, this approach is also sensitive to noise. In order to improve the noise robustness, the size of the window is changed adaptively, as discussed in the next section.

## 4.2 Noise robustness

The features extracted by using LBP are easily affected if any slight change occurs in the intensity value of the central pixel value due to noise. In order to have the robustness in obtaining the features by LBP, a new window for a given window size  $N \times N$  is defined as follows: The original central pixel value of the window is replaced with the mean value of the pixels in the window, because the original central pixel value is approximately equal to the mean value of the pixels in the given window. Next, the intensity values of neighbours of given central pixel are modified by averaging the intensity values along radial direction [42]. This procedure of averaging along radial directions is explained in Fig. 4 for  $N = 5$ . However, the size of the window also affects the recognition rate. If the window size is varied according to the variance of the intensity values within the window, then the proposed approach improve the performance of FER system.

## 4.3 Adaptive window

In this section, the method of selection of window size is discussed. The modification of window for feature extraction by averaging along radial direction achieves the noise robustness. However, the selection of window size for averaging the pixel values in radial direction influences the feature description. Hence, the appropriate selection of window size is necessary before feature extraction. The characteristics of the feature vector changes if the window size is varied. The variance of the pixel intensities within the window is less, and then it is considered as smooth region, else it is considered as non-smooth region. The size of the window is high for smooth region, whereas the size of the window is less for non-smooth region. This has motivated to vary the size of the window according to variance of intensity values within the window. If  $V_1$ ,  $V_2$  and  $V_3$  are the variances of windows with



$$g_c = \text{mean}(P7, P8, P9, P12, P13, P14, P17, P18, P19)$$

$$g_1 = \frac{P1 + P7 + P13}{3}, \quad g_2 = \frac{P3 + P8 + P13}{3}$$

$$g_3 = \frac{P5 + P9 + P13}{3}, \quad g_4 = \frac{P11 + P12 + P13}{3}$$

$$g_5 = \frac{P15 + P14 + P13}{3}, \quad g_6 = \frac{P21 + P17 + P13}{3}$$

$$g_7 = \frac{P23 + P18 + P13}{3}, \quad g_8 = \frac{P25 + P19 + P13}{3}$$

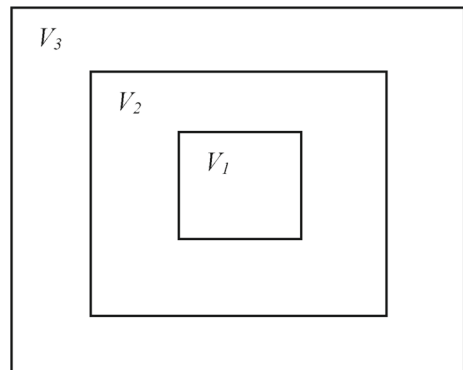
Fig. 4 Modified neighborhood for noise robustness

different sizes as shown in Fig. 5, then the size of the window is selected based on the following steps:

- (i) If  $V_1 < V_2$ , then window size is increased because of the smoothness of the region. Otherwise, original window size is retained.
- (ii) If  $V_2 < V_3$ , then window size is further increased, otherwise same window size is retained.

After selection of window size using above steps, the intensity values of neighbours around a central pixel are modified by the procedure shown in Fig. 4. After modifying the intensity values of a given window with any size  $N \times N$  to  $3 \times 3$  window, the features are extracted by using (8). The schematic representation of the proposed method for feature extraction shown in Fig. 6. It is observed that the features extracted on the whole face image

Fig. 5 Adaptive window approach:  $V_1, V_2$  and  $V_3$  are variances





contains only macro level information [19]. In order to obtain the micro level information, each face image is divided into sub regions and features are extracted from each sub region [15]. Finally, histograms from each sub image are concatenated to get the final feature of the face image [31]. Since the proposed method considers both averaging in radial directions and adaptive window, the performance is robust in terms of recognition rate in noisy images.

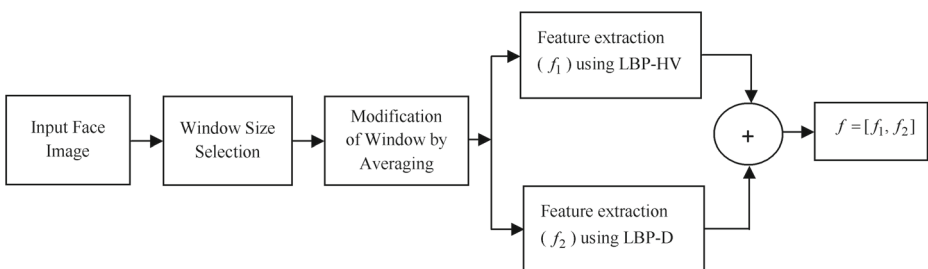
The novelty of proposed method compared to LBP is that it has three advantages. First, the traditional LBP uses eight neighbors which results in feature length of 256 whereas our method has feature length of 32 only. For example, consider an image with dimension  $128 \times 128$ . Suppose if it is divided into  $8 \times 8$  sub regions, then extracted feature has length of 16384 ( $8 \times 8 \times 256 = 16384$ ) whereas our approach has a dimension of 2048 ( $8 \times 8 \times 32 = 2048$ ) only. Such a reduction in feature characteristic dimension is useful at the stage of classification. Second, the formation of window for feature extraction by averaging pixels in radial directions makes proposed method more robust to the noise, but the traditional LBP is highly sensitive to noise. Third, the concept of adaptive window in the selection of window size for averaging in radial direction further enhances the performance of FER system.

## 5 Experimental setup

In this section, a brief discussion on databases used and the procedure for performance evaluation of the proposed method are explained. Japanese Female Facial Expression (JAFFE) database [37] and Cohn-Kanade (CK) database [29] are popularly used by the researchers in the area of FER. Each database contains seven basic emotions: neutral, anger, disgust, fear, happy, sad, and surprise.

### 5.1 JAFFE database

This database contains the images of 10 Japanese females and each person was instructed to display the six basic prototypic expressions starting from the neutral expression. It has a total of 213 images (30 anger, 29 disgust, 32 fear, 31 joy, 31 sadness, 30 surprise, and 30 neutral) with each image has a dimension of  $256 \times 256$  pixels. Some of the sample images from JAFFE database are shown in Fig. 7.



**Fig. 6** Schematic representation of the proposed method for feature extraction



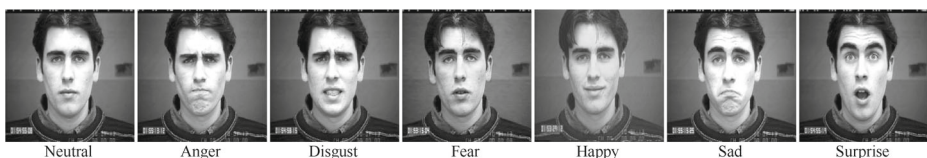
**Fig. 7** Sample face expression images from JAFFE database

## 5.2 CK database

This database contains the images of 100 university students aged from 18 to 30 years, of which 65% were female, 15% were African-American and 3% were Asian or Latino. Students were instructed to display the facial expressions starting from neutral expression to target expression; images were digitized to a dimension of  $640 \times 480$  or  $490$ . In this work, first neutral frame and last three peak frames of each sequence are selected in the database. In total, 1290 images are used (108 anger, 117 disgust, 105 fear, 276 joy, 129 sadness, 225 surprise, and 330 neutral). Some of the sample images from CK database are shown in Fig. 8.

## 5.3 Testing procedure of the proposed method

The selected images from the databases are cropped to a size of  $150 \times 110$  as done in [48]. Further, each image is divided into  $7 \times 6$  sub regions to get the micro-level information from the face. After extraction of features, the Principal Component Analysis (PCA) is applied for dimensionality reduction. The recognition rate is calculated to evaluate the performance of the FER system by using proposed method. The recognition rate is calculated by dividing the number of correct predictions by the total number of predictions. The 10-fold cross validation testing scheme is used to calculate the recognition rate. In 10-fold cross validation, the dataset is roughly divided into ten equal groups. Nine groups are used for training and one group is used for testing. This process is repeated 10 times and the recognition rates are averaged to obtain the final recognition rate. The confusion matrix is also calculated in order to obtain the recognition rate of individual facial expression. The confusion matrix in machine learning describes the performance of a given classifier on testing data for which true labels are known. In confusion matrix, the rows correspond to the observed classes and the columns correspond to the predicted labels. The diagonal elements in the confusion matrix represent the correctly classified classes and off-diagonal elements give incorrectly classified classes. Finally, for classification, the support vector machine (SVM) is used because of its popularity in most of the pattern recognition problems [12]. SVM is a supervised machine learning algorithm that can be used for binary classification. It performs classification by finding a linear separating hyperplane that maximizes the margin



**Fig. 8** Sample face expression images from CK database

between the classes in high dimensional space. SVM is a binary classifier, and the multi-class classification is obtained by using one-against-rest technique [34]. All experiments are done by using MATLAB software. The classification learner app of MATLAB is used for SVM implementation and quadratic kernel is selected.

## 6 Results and discussion

In this section, results on the JAFFE database and the CK database are discussed for the both 6-class and 7-class FER. In the first set of experiments, the proposed method is evaluated for different sizes of the window  $N \times N$  and adaptive window on the JAFFE and the CK databases. The values of  $N$  are selected as 3, 5, and 7 respectively, for the size of the window. Experimental results are shown in Tables 1 and 2 for the JAFFE and the CK databases. It can be seen from Tables 1 and 2, the adaptive window approach gives the highest recognition rate of 92.9% and 88.3% for 6-class and 7-class recognition on the JAFFE database. Similarly, this approach achieves recognition rate of 96% and 93.9% for 6-class and 7-class recognition on the CK database. The recognition rate for the adaptive window is higher because of the size of the window is changed according to the variance of the pixel intensities within the window.

In the second set of experiments, in order to know the performance of the proposed method on the individual facial expressions, the confusion matrices for 6-class and 7-class FER on the JAFFE database are calculated. The confusion matrices are given for both 6-class and 7-class in Tables 3 and 4 respectively. It is observed that, happy and surprise are recognised with higher recognition rates of 100% and 96.7%, while recognition rate for anger is lower and easily confused with the other facial expressions. As neutral expression is included in 7-class recognition, the recognition rate is reduced compared to 6-class recognition rate because some of the facial expressions are confused with the neutral expression.

The same experimental procedure is performed on the CK database, and confusion matrices for 6-class and 7-class recognition are given in Tables 5 and 6 respectively. The overall performance in the CK database is more compared to the JAFFE database because of the large number of subjects is available for facial expressions in the CK database. Another reason for lower recognition rate on the JAFFE database is that some facial expressions in this database had been labelled incorrectly and these images influence the recognition rate.

The proposed method is compared with the traditional methods LBP [48], LTP [51], LGC-HD [53], IGLTP [21], HOG [16], Wavelet + HOG [40] and few deep learning methods: Deep CNN [50], CNN [28], DNN [28], Deep CNN [36], DNN [33] and WMDNN [54]. The simulation results on the JAFFE and CK databases are given in Tables 7 and 8. It is observed that the proposed method outperformed with a recognition rate of 92.9% for 6-

**Table 1** Recognition rate(%) on JAFFE database

Window size	6-class Recognition	7-class Recognition
3 x 3	90.2	86.4
5 x 5	91.3	87.5
7 x 7	91.2	87.9
Adaptive window	92.9	88.3

**Table 2** Recognition rate(%) on CK database

Window size	6-class Recognition	7-class Recognition
3 x 3	95.1	92.3
5 x 5	95.3	93.1
7 x 7	94.7	92.9
Adaptive window	96	93.9

**Table 3** Confusion matrix of 6-class FER on JAFFE database

	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	<b>86.6</b>	6.7	0	0	6.7	0
Disgust	6.8	<b>89.7</b>	0	0	3.5	0
Fear	0	3.1	<b>93.8</b>	0	3.1	0
Happy	0	0	0	<b>100</b>	0	0
Sad	0	0	6.5	3.2	<b>90.3</b>	0
Surprise	0	0	0	3.3	0	<b>96.7</b>

The bold entries signifies the recognition rate of individual facial expression

**Table 4** Confusion matrix of 7-class FER on JAFFE database

	Anger	Disgust	Fear	Happy	Sad	Surprise	Neutral
Anger	<b>84.2</b>	8.0	0	0	3.3	0	4.5
Disgust	3.5	<b>79.3</b>	13.7	0	0	3.5	0
Fear	0	6.2	<b>87.6</b>	0	6.2	0	0
Happy	0	0	0	<b>93.6</b>	3.2	0	3.2
Sad	0	0	6.5	3.2	<b>87.1</b>	0	3.2
Surprise	0	0	3.3	3.3	0	<b>93.4</b>	0
Neutral	2.5	3.4	0	0	1.5	0	<b>92.6</b>

The bold entries signifies the recognition rate of individual facial expression

**Table 5** Confusion matrix of 6-class FER on CK database

	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	<b>91.6</b>	2.8	0	0.9	4.7	0
Disgust	1.7	<b>94</b>	0	0	4.3	0
Fear	0	0	<b>96.2</b>	2.9	0	0.9
Happy	0	0	0	<b>100</b>	0	0
Sad	1.6	0.7	0	0	<b>96.9</b>	0.7
Surprise	0	0	0.4	0	0.4	<b>99.2</b>

The bold entries signifies the recognition rate of individual facial expression

**Table 6** Confusion matrix of 7-class FER on CK database

	Anger	Disgust	Fear	Happy	Sad	Surprise	Neutral
Anger	<b>85.3</b>	2.8	0	0.9	1.8	0	9.2
Disgust	0.9	<b>92.4</b>	0	0	2.5	0	4.2
Fear	0	0	<b>95.2</b>	3.9	0	0	0.9
Happy	0	0	0.3	<b>97.5</b>	0	0	2.2
Sad	1.6	0	0	0	<b>88.4</b>	0.7	9.3
Surprise	0	0.4	0	0	0.4	<b>96.5</b>	2.7
Neutral	0.6	0	0	0	0	0	<b>99.4</b>

The bold entries signifies the recognition rate of individual facial expression

**Table 7** Comparison of recognition rate(%) on JAFFE database for various methods

Method	6-class Recognition	7-class Recognition
LBP [48]	84.7	81.2
LTP [51]	86.9	82.6
LGC-HD [53]	89.2	86.6
IGLTP [21]	89.9	86.9
HOG [16]	–	76.19
Wavelet+HOG [40]	–	71.43
Deep CNN [50]	–	76.7
DNN [33]	91.27	–
WMDNN [54]	92.21	–
Proposed	92.9	88.3

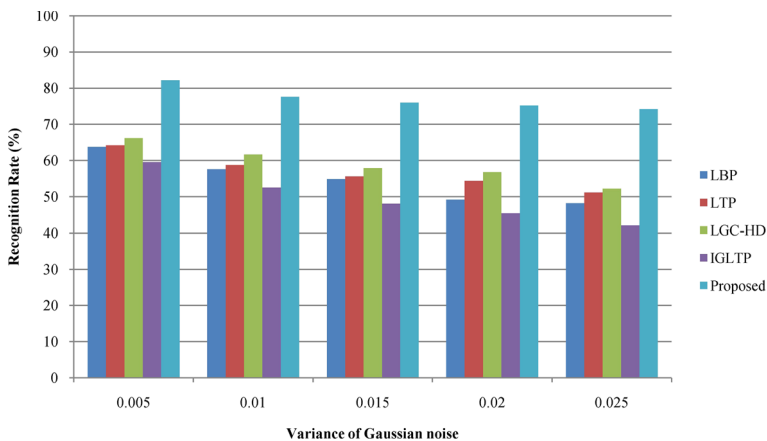
**Table 8** Comparison of recognition rate(%) on CK database for various methods

Method	6-class Recognition	7-class Recognition
LBP [48]	88.9	87.4
LTP [51]	91.9	88.4
LGC-HD [53]	92.1	89.2
IGLTP [21]	93.6	90.6
Wavelet+HOG [40]	90.0	–
Deep CNN [50]	–	80.3
CNN [28]	–	86.5
DNN [28]	–	72.7
Deep CNN [36]	–	92.81
DNN [33]	96.46	–
WMDNN [54]	97.02	–
Proposed	96	93.9

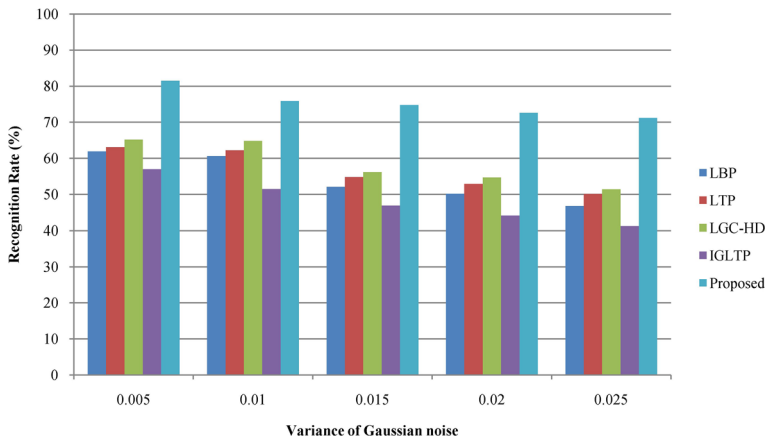
class and 88.3% for 7-class compared to the existing methods on JAFFE database. Similarly the results on the CK database for 6-class and 7-class are given in Table 8. As it can be seen from Table 8, the recognition rate of the proposed method is 96% for 6-class and 93.9% for 7-class. It is observed that Deep learning methods: DNN [33] and WMDN [54] achieves better performance with recognition rates 96.4% and 97.02% on CK database but comparable with the proposed method which is 96%. However, on JAFFE database, the proposed method outperforms than DNN [33] and WMDN [54].

Finally, in order to verify the robustness of the proposed method against noise, the Gaussian noise is added with zero mean and different variances equal to 0.05, 0.01, 0.015, 0.02, 0.025 to the images of the JAFFE and the CK databases. For comparison, The LBP and its variants: LBP, LTP, LGC-HD, and IGLTP are considered because these methods are highly effected by noise. The simulation results are shown in Figs. 9 and 10 for 6-class and 7-class on the JAFFE database. It is observed that IGLTP has poor performance in the presence of noise, but it outperformed without noise compared to the LBP, LTP, LGC-HD methods. But, the proposed method gives good recognition rate compared to all these methods. Similarly, the simulation results on the CK database are shown in Figs. 11 and 12 for 6-class and 7-class respectively. The simulation results confirm that the proposed method gives good recognition rate on the CK database when compared to the LBP and its variants. The experimental results validate that proposed approach in calculating the LBP with adaptive window is a powerful approach against noise and noise free conditions.

Further, the proposed method is evaluated on two other publicly available databases: Facial Expression Research Group (FERG) database [3] and FEI face database [52]. The FERG Database has 55767 annotated face images of six stylized characters namely: Ray, Malcolm, Jules, Bonnie, Mery, and Aia. Each character has seven types of universal facial expressions namely: neutral, anger, disgust, fear, happy, sad, and surprise with original size of  $768 \times 768$  pixels. Six sample images from this database are shown in Fig. 13. The FEI face database is a Brazilian face database contains 14 images for each of 200 persons with original size of  $640 \times 480$  pixels. All images are obtained from students and staff in between 19 and 40 years old. Each person is posed two basic emotions: neutral and happy. Some sample images from FEI face database are shown in Fig. 14.

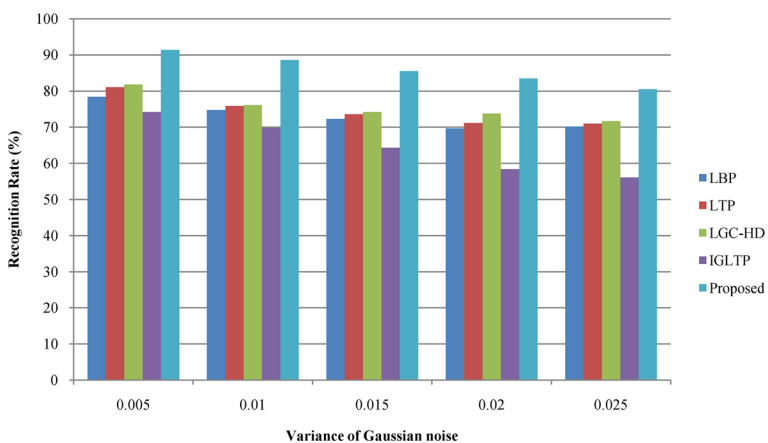


**Fig. 9** Recognition Rate on JAFFE database for 6-class under different levels of Gaussian noise



**Fig. 10** Recognition Rate on JAFFE database for 7-class under different levels of Gaussian noise

In our experiments, 15 images from each character with seven basic emotions from FERF database are considered. All color images from this database are converted into gray scale images. Similarly, all 200 images with frontal pose from FEI face database are considered for 2-class expression recognition. The experimental procedure is same as conducted for JAFFE and CK databases. The average recognition rates of the proposed method and some of the previous methods are provided in Table 9. It is noticed that the average recognition rates on FERF data base and FEI face database are 96.7% and 98.9% respectively which outperforms compared to the existing methods. It is also observed that best recognition rate FEI face data base is obtained because it has only two emotions whereas other databases have all seven basic emotions. In summary, it is confirmed that the proposed method achieves higher recognition rate on JAFFE, CK, FERF and FEI face databases compared to the existing methods.



**Fig. 11** Recognition Rate on CK database for 6-class under different levels of Gaussian noise

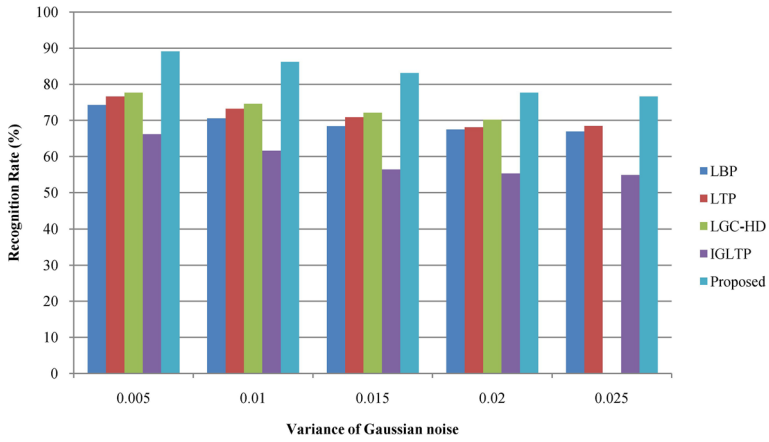


Fig. 12 Recognition Rate on CK database for 7-class under different levels of Gaussian noise

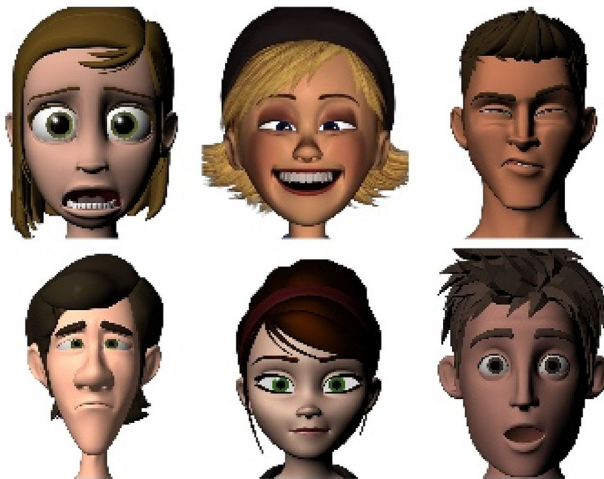


Fig. 13 Sample images from FERF database



Fig. 14 Sample Images from FEI face database



**Table 9** Performance of proposed method on other databases

Database	Method	Recognition Rate(%)
FERG	LBP [48]	89.6
	LTP [51]	90.5
	LGC-HD [53]	92.3
	IGLTP [21]	95.1
	DeepExpr [3]	89.1
	Proposed	96.7
FEI	LBP [48]	93.7
	LTP [51]	94.9
	LGC-HD [53]	96.4
	IGLTP [21]	97.8
	CLTP [55]	96.3
	Proposed	98.9

## 7 Conclusion

The performance of any FER system strongly depends on the efficient feature extraction. This paper presents a novel approach for feature extraction. The major contribution of this work is to improve the LBP for feature extraction. The highlights of the proposed method are: it is a powerful texture descriptor for FER; it reduces the feature dimension which further improves the classification; and finally it is more robust against noise. The features are extracted by applying LBP separately to the 4-neighbors and diagonal neighbors. This approach has the advantage of less size in dimension of the feature compared to traditional LBP. The effect of noise in the feature extraction is reduced by using new neighbourhood which is obtained by averaging in the radial directions and the adaptive window concept. The combination of adaptive window and averaging in radial direction make the feature as more affective in texture description. The classification is performed by using Support Vector Machine. Experimental results on the JAFFE, CK, FERG and FEI face databases confirm that the superiority of the proposed method in recognition rate compared to the existing methods. In future work, the use of deep learning concepts can be explored for classification and to improve recognition rate. Further, the concepts of SOD using deep learning can be applied for FER system.

**Acknowledgments** The corresponding author acknowledges the research colleagues Dr. G. Sridevi, Professor of ECE, Aditya Engineering College, Surampalem, and Dr. B. Chandra Mohan, Professor of ECE, Bapatla Engineering College, Bapatla, for their valuable advices and suggestions during the execution of this work.

## Compliance with Ethical Standards

**Conflict of interests** The authors declare that they have no conflict of interest.

## References

1. Ahonen T, Hadid A, Pietikainen M (2006) Face description with local binary patterns: application to face recognition. *IEEE Trans Pattern Anal Mach Intell* 28(12):2037–2041

2. An Q, Han Y, Li J, Lu S (2018) Human-computer interaction nursing system and related algorithms for severely paralyzed patients. In: 2018 15th International conference on control, automation, robotics and vision (ICARCV). IEEE, pp 1929–1935
3. Aneja D, Colburn A, Faigin G, Shapiro L, Mones B (2016) Modeling stylized character expressions via deep learning. In: Asian conference on computer vision. Springer, pp 136–153
4. Awad AI, Hassaballah M (2016) Image feature detectors and descriptors. Studies in Computational Intelligence Springer International Publishing, Cham
5. Bashyal S, Venayagamoorthy GK (2008) Recognition of facial expressions using gabor wavelets and learning vector quantization. *Eng Appl Artif Intell* 21(7):1056–1064
6. Bellamkonda S, Gopalan N (2018) Facial expression recognition using kirsch edge detection, lbp and gabor wavelets. In: 2018 Second international conference on intelligent computing and control systems (ICICCS). IEEE, pp 1457–1461
7. Bi H, Li N, Guan H, Lu D, Yang L (2019) A multi-scale conditional generative adversarial network for face sketch synthesis. In: 2019 IEEE international conference on image processing (ICIP). IEEE, pp 3876–3880
8. Chao WL, Ding JJ, Liu JZ (2015) Facial expression recognition based on improved local binary pattern and class-regularized locality preserving projection. *Signal Process* 117:1–10
9. Chen A, Xing H, Wang F (2020) A facial expression recognition method using deep convolutional neural networks based on edge computing. *IEEE Access* 8:49741–49751
10. Chengeta K, Viriri S (2019) A review of local, holistic and deep learning approaches in facial expressions recognition. In: 2019 Conference on information communications technology and society (ICTAS). IEEE, pp 1–7
11. Cho M, Kim T, Kim IJ, Lee S (2020) Relational deep feature learning for heterogeneous face recognition. [arXiv:200300697](https://arxiv.org/abs/200300697)
12. Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20(3):273–297
13. Dan Z, Chen Y, Yang Z, Wu G (2014) An improved local binary pattern for texture classification. *Optik* 125(20):6320–6324
14. Donato G, Bartlett MS, Hager JC, Ekman P, Sejnowski TJ (1999) Classifying facial actions. *IEEE Trans Pattern Anal Mach Intell* 21(10):974–989
15. Ekweariri AN, Yurtkan K (2017) Facial expression recognition using enhanced local binary patterns. In: 2017 9th International conference on computational intelligence and communication networks (CICN). IEEE, pp 43–47
16. Eng S, Ali H, Cheah A, Chong Y (2019) Facial expression recognition in jaffe and kdef datasets using histogram of oriented gradients and support vector machine. In: IOP Conference series: materials science and engineering, vol 705. IOP Publishing, p 012031
17. Fan DP, Cheng MM, Liu JJ, Gao SH, Hou Q, Borji A (2018) Salient objects in clutter: Bringing salient object detection to the foreground. In: Proceedings of the European conference on computer vision (ECCV), pp 186–202
18. Fan DP, Zhang S, Wu YH, Liu Y, Cheng MM, Ren B, Rosin PL, Ji R (2019) Scoot: A perceptual metric for facial sketches. In: Proceedings of the IEEE international conference on computer vision, pp 5612–5622
19. Farajzadeh N, Hashemzadeh M (2018) Exemplar-based facial expression recognition. *Inf Sci* 460:318–330
20. Hassaballah M, Awad AI (2016) Detection and description of image features: an introduction. In: Image feature detectors and descriptors. Springer, pp 1–8
21. Holder RP, Tapamo JR (2017) Improved gradient local ternary patterns for facial expression recognition. *EURASIP J Image Video Process* 2017(1):42
22. Huang D, Shan C, Ardabilian M, Wang Y, Chen L (2011) Local binary patterns and its application to facial image analysis: a survey. *IEEE Trans Syst Man Cybern Part C App Rev* 41(6):765–781
23. Huang Z, Song G, Zhao Y, Han J, Zhao X (2018) Smile recognition based on support vector machine and local binary pattern. In: 2018 IEEE 8th Annual international conference on cyber technology in automation, control, and intelligent systems (CYBER). IEEE, pp 938–942
24. Huang Y, Wang Y, Tai Y, Liu X, Shen P, Li S, Li J, Huang F (2020) Curricularface: adaptive curriculum learning loss for deep face recognition. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 5901–5910
25. Jabid T, Kabir MH, Chae O (2010) Local directional pattern (ldp) for face recognition. In: 2010 Digest of technical papers international conference on consumer electronics (ICCE). IEEE, pp 329–330
26. Jabid T, Kabir MH, Chae O (2010) Robust facial expression recognition based on local directional pattern. *ETRI J* 32(5):784–794

27. Jaiswal S, Nandi G (2019) Robust real-time emotion detection system using cnn architecture. *Neural Comput & Applic*, 1–10
28. Jung H, Lee S, Park S, Kim B, Kim J, Lee I, Ahn C (2015) Development of deep learning-based facial expression recognition system. In: 2015 21st Korea-Japan joint workshop on frontiers of computer vision (FCV). IEEE, pp 1–4
29. Kanade T, Cohn JF, Tian Y (2000) Comprehensive database for facial expression analysis. In: Proceedings Fourth IEEE international conference on automatic face and gesture recognition (Cat. No. PR00580). IEEE, pp 46–53
30. Kaplan K, Kaya Y, Kuncan M, Minaz MR, Ertunç HM (2020) An improved feature extraction method using texture analysis with lbp for bearing fault diagnosis. *Appl Soft Comput* 87:106019
31. Kaushik MS, Kandali AB (2017) Recognition of facial expressions extracting salient features using local binary patterns and histogram of oriented gradients. In: 2017 International conference on energy, communication, data analytics and soft computing (ICECDS). IEEE, pp 1201–1205
32. Khan RA, Meyer A, Konik H, Bouakaz S (2013) Framework for reliable, real-time facial expression recognition for low resolution images. *Pattern Recogn Lett* 34(10):1159–1168
33. Kim JH, Kim BG, Roy PP, Jeong DM (2019) Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. *IEEE Access* 7:41273–41285
34. Lekdioui K, Messoussi R, Ruichek Y, Chaabi Y, Touahni R (2017) Facial decomposition for expression recognition using texture/shape descriptors and svm classifier. *Signal Process Image Commun* 58:300–312
35. Li J, Zhang D, Zhang J, Zhang J, Li T, Xia Y, Yan Q, Xun L (2017) Facial expression recognition with faster r-cnn. *Procedia Comput Sci* 107:135–140
36. Liliana D (2019) Emotion recognition from facial expression using deep convolutional neural network. In: *Journal of physics: conference series*, vol 1193. IOP Publishing, p 012004
37. Lyons M, Akamatsu S, Kamachi M, Gyoba J (1998) Coding facial expressions with gabor wavelets. In: Proceedings Third IEEE international conference on automatic face and gesture recognition. IEEE, pp 200–205
38. Mehta R, Egiazarian K (2016) Dominant rotated local binary patterns (drlbp) for texture classification. *Pattern Recogn Lett* 71:16–22
39. Minaee S, Abdolrashidi A (2019) Deep-emotion: Facial expression recognition using attentional convolutional network. [arXiv:1902.01019](https://arxiv.org/abs/1902.01019)
40. Nigam S, Singh R, Misra A (2018) Efficient facial expression recognition using histogram of oriented gradients in wavelet domain. *Multimed Tools Appl* 77(21):28725–28747
41. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Mach Intell* 24(7):971–987
42. Pan Z, Wu X, Li Z, Zhou Z (2017) Local adaptive binary patterns using diamond sampling structure for texture classification. *IEEE Signal Process Lett* 24(6):828–832
43. Patil M, Veni S (2019) Driver emotion recognition for enhancement of human machine interface in vehicles. In: 2019 International conference on communication and signal processing (ICCSP). IEEE, pp 0420–0424
44. Perez-Gaspar LA, Caballero-Morales SO, Trujillo-Romero F (2016) Multimodal emotion recognition with evolutionary computation for human-robot interaction. *Expert Syst Appl* 66:42–61
45. Pitaloka DA, Wulandari A, Basaruddin T, Liliana DY (2017) Enhancing cnn with preprocessing stage in automatic emotion recognition. *Procedia Comput Sci* 116:523–529
46. Roy SD, Bhowmik MK, Saha P, Ghosh AK (2016) An approach for automatic pain detection through facial expression. *Procedia Comput Sci* 84:99–106
47. Salahat E, Qasaimeh M (2017) Recent advances in features extraction and description algorithms: a comprehensive survey. In: 2017 IEEE international conference on industrial technology (ICIT). IEEE, pp 1059–1063
48. Shan C, Gong S, McOwan PW (2005) Robust facial expression recognition using local binary patterns. In: *IEEE International conference on image processing 2005*, vol 2. IEEE, pp II–370
49. Shan C, Gong S, McOwan PW (2009) Facial expression recognition based on local binary patterns: a comprehensive study. *Image Vision Comput* 27(6):803–816
50. Shan K, Guo J, You W, Lu D, Bie R (2017) Automatic facial expression recognition based on a deep convolutional-neural-network structure. In: 2017 IEEE 15th international conference on software engineering research, management and applications (SERA). IEEE, pp 123–128
51. Tan X, Triggs B (2010) Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans Image Process* 19(6):1635–1650
52. Thomaz CE, Giraldi GA (2010) A new ranking method for principal components analysis and its application to face image analysis. *Image Vision Comput* 28(6):902–913

53. Tong Y, Chen R, Cheng Y (2014) Facial expression recognition algorithm using lgc based on horizontal and diagonal prior principle. *Optik* 125(16):4186–4189
54. Yang B, Cao J, Ni R, Zhang Y (2017) Facial expression recognition using weighted mixture deep neural network based on double-channel facial images. *IEEE Access* 6:4630–4640
55. Yee SY, Rassem TH, Mohammed MF, Makbol NM (2019) Performance evaluation of completed local ternary pattern (cltp) for face image recognition. *Perform Eval*, 10(4)
56. Zhang J, Fan DP, Dai Y, Anwar S, Saleh FS, Zhang T, Barnes N (2020) Uc-net: uncertainty inspired rgb-d saliency detection via conditional variational autoencoders. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp 8582–8591
57. Zhang Y, Hua C (2015) Driver fatigue recognition based on facial expression analysis using local binary patterns. *Optik* 126(23):4501–4505
58. Zhang J, Yu X, Li A, Song P, Liu B, Dai Y (2020) Weakly-supervised salient object detection via scribble annotations. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp 12546–12555
59. Zhao JX, Liu JJ, Fan DP, Cao Y, Yang J, Cheng MM (2019) Egnet: Edge guidance network for salient object detection. In: *Proceedings of the IEEE international conference on computer vision*, pp 8779–8788

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.