A Texture Features-Based Robust Facial Expression Recognition

Jayati Krishna Goswami, GLA University, India
Sunita Jalal, G. B. Pant University of Agriculture and Technology, India
Chetan Singh Negi, G. B. Pant University of Agriculture and Technology, India
Anand Singh Jalal, GLA University, India

https://orcid.org/0000-0002-7469-6608

ABSTRACT

Facial expression plays an important role in communicating emotions. In this paper, a robust method for recognizing facial expressions is proposed using the combination of appearance features. Traditionally, appearance features mainly divide any face image into regular matrices for the computation of facial expression recognition. However, in this paper, the authors have computed appearance features in specific regions by extracting facial components such as eyes, nose, mouth, and forehead. The proposed approach mainly has five stages to detect facial expression viz. face detection and regions of interest extraction, feature extraction, pattern analysis using a local descriptor, the fusion of appearance features, and finally, classification using a multiclass support vector machine (MSVM). Results of the proposed method are compared with the earlier holistic representations for recognizing facial expressions, and it is found that the proposed method outperforms state-of-the-art methods.

KEYWORDS

Emotion Recognition, Face Emotion, Feature Extraction, LDOP

1. INTRODUCTION

Over the last decades, researchers performed various analysis related to facial expression recognition in the field of computer vision and pattern recognition. Generally, there are three types of communication - verbal, non-verbal and para verbal. In general, in any conversation, verbal, non-verbal and para verbal contribute 7%, 55%, and 38%, respectively. Facial expression plays an essential role in non-verbal communication. Face expression also plays a significant role in between human to human or human to machine communication. Muscle movement under facial skin causes facial expression, and these movements change according to internal or external emotions. Eye contact establishes participation, conversations and generates a bond with others.

In the early 70s, a famous psychologist Paul Ekman (Ekman & Friesen, 1978) and his associates performed research on facial expressions and proposed six primary facial expressions. Facial expressions comprise the laugh, sad, anger, disgust, surprise, and fear. The face with the curve-shaped eye expresses happiness. In the sad expression, eyebrows’ inner corners are drawn in and triangulation
of skin underneath the eyebrows and scowl. An angry face is related to the obnoxious and bothersome states. The anger facial expression is expressed with compressed eyebrows, thin and drawn eyelids. In a disgusted face, the upper eyelid is lifted up, and eyebrows are pull-down and wrinkled nose. In a surprised face, the skin below the eyebrow is drawn, widening of eyes and mouth gaping which can be recognized easily. The fear expression can be identified on a face having a wrinkled forehead between the eyebrows and tensed lips and opened mouth (Revina and Emmanuel, 2018).

Fig. 1 demonstrates the general Facial Expression Recognition (FER) framework. In general, a FER system comprises two stages - extraction of features and classification of the facial expression. Extraction of the features can be performed using two methods – geometric-based methods and appearance-based methods (Dols & Russell, 2017) (Tian et al., 2011) (Chanti et al., 2017). In the geometric-based method, facial landmarks like eyes, forehead, nose, and mouth are located to describe facial geometry. Although, facial analysis applications (Viola & Jones, 2004) (Bartlett et al., 2002) (Bartlett et al., 2005) based on the geometric feature-based methods yield more approving outcomes, however, it has some issues in computing in various situations.

On the contrary, in the appearance-based method, filters are applied over an image holistically or on the regions specifically to discriminate variations in facial expressions (Liang et al., 2016). Even though Local Binary Pattern (reference) is adopted frequently for extracting texture-based features and captures the local region characteristics from images in which facial expression recognition is performed. However, existing facial expression recognition methods subject to LBP is not capable of solving the issues of local illuminations that make recognition task difficult due to uncertainty in some facial regions.

The block-level processing of the face images plays a key role in finding details and boundary by placing the center on landmarks. Also, the dimensionality of feature extraction is increased, and it is inconvenient to specify every characteristic of image texture. Guo et al. (2017) introduced a recognition technique for facial expression called “K-ELBP” in which K-L transform (KLT) is merged with Extended Local Binary Pattern (ELBP). The feature values are extracted using ELBP by applying it on expression images. The feature vector is computed by reducing the dimensionality of the feature matrix using a covariance matrix transform. Finally, for the classification of the facial expression, a Support Vector Machine (SVM) is utilized.

Further, to resolve various other issues of LBP, Arshid et al. (2018), introduced a feature called MSBP (Multi-Stage Binary pattern) in which a binary code is produced for each comparison at multiple steps.

Figure 1. Generic framework of FER
stages against neighbouring pixel. LBP drops a few important texture information by recording only a sign difference. Gradient difference and sign difference is combined to form MSBP which accumulate variations in the region encompassing eyebrows, eyes, mouth, etc.

There are two different scenarios in the proposed approach: holistic and zone-based. Also, a method called double delta-LBP is proposed by Shen et al. 2018 in which both detail and contour face information are utilized by applying two delta-LBP’s in the feature extraction process to recognize facial expression. Histograms of oriented gradients (HOGs) method has been widely used descriptor in which orientation density pattern and gradient of the edge are used in preserving the local information. In (Nazir et al., 2018) HOG features have been examined. For the specification of illumination and orientation invariant facial expressions, the transformation of HOG features into the frequency domain is done to make descriptor more suitable. For the transformation of the features in the frequency domain, Discrete cosine transform (DCT) is used, and the most important discriminant features are extracted. KNN classifier (k-nearest neighbour) is employed for the facial expression is classified.

Most of the existing work utilized global features for facial expression recognition. However, the global features are failed to deal with local variations among the different facial components.

Problem statement: Appearance features have been computed only in specific regions by extracting facial components in which the discrimination rate is high.

The contribution of the proposed work is summarized as given below:

1. Features extracted using Histogram of Oriented Gradients (HOGs) are fused with Local Directional Order pattern by which discrimination and robustness are improved and able to give better result in recognizing facial expressions.
2. The method has robustness against orientations and illumination.
3. This approach includes encoding of the broader neighbourhood for the discrimination of inter-class likeness and to attain the robustness for the intra-class variations.

The organization of the paper is as follows: The related work is discussed in Section 2. Section 3 describes the proposed approach to facial expression recognition. Result and discussion are performed in Section 4 and lastly, Section 5 elaborates the conclusion.

2. RELATED WORK

Generally, a facial expression recognition scheme has three phases: Face detection, Extraction of Features, and Expression Classification. Three main challenges related to the FER system are needed to be considered. The first foremost challenge is to select features which, according to this study, should be picked carefully for better outcomes (Subramanyam & Fegade, 2013). The next challenge is about classification: It all depends upon the choice of the dataset which will improve the accuracy. The machine-learning algorithm can be utilized for achieving greater accuracy and a distinctive database can be used for the experimentation.

Another key problem is associated with multiple backgrounds and unmonitored lightening conditions. Moreover, this paper also discussed two other problems with the orientation and scale of the images. The three-stage local directional based approach suggested by Faraji & Qi (2016). For handling illumination. Initially, the harmonic filter is utilized and then, facial features are enhanced using LFD and finally, illumination invariant features are represented using directions and magnitudes of edge responses. In the last decade, various feature extraction methods have been introduced for the recognition of facial expression. Several studies based on facial expression recognition deal with challenges in this domain and present reasonable solutions (Corneanu et al., 2016) (Poria et al.,
feature learning is effective for all features extraction associated with facial expression. Ekman and Friesen (1978) proposed a geometric-feature based technique which is very popular among approaches for facial action coding system (FACS).

A set of action units (AU) is used for the identification of Facial expressions and these action units are the physical movement of a particular facial muscle beneath the skin. Guo & Dyer (2003) proposed a geometric based facial feature representation method that positioned facial points manually. Likewise, an approach is introduced in Valster & Pantic (2006) in which feature extraction and training of classifier are done using linear programming. Similar to this, facial expression analysis systems based on tracked point data have been proposed in Valstar et al. (2005, 2006, 2010).

Nigam et al. (2018) gave a new technique which is based on Discrete Wavelet Transform and Histogram of oriented gradients are introduced. This technique named a W_HOG feature (Nigam et al., 2018). Discrete wavelet transform (DWT) is applied to spatial domain features to transform into the frequency domain. The Histogram of Gradients (HOG) is applied along with DWT for the feature extraction and this combined feature is called a W_HOG feature. Multiclass Support Vector Machine (SVM) classifier used, for the classification of facial expression, in which proposed W_HOG feature is taken into.

Yang et al. (2018) gave a fusion method for facial expression recognition. In this method, Gabor wavelets are utilized for the representation of an image, which renders well regional spatial features and orientation selectivity of image textures. These features have the robustness to intra-class variance. Moreover, the dimensionality of the Gabor feature vector is decreased and Multi-Orientation Symmetric Local Graph Structure (MSLGS) is introduced to achieve better discrimination and less computation cost during the feature extraction process. The orientation of the original WLD (Weber Local Descriptor) is extended by including more gradient direction for obtaining the spatial structure information.

Ghimire et al. (2017) used the combination of both appearance-based and geometric-based features for facial expression recognition. Face region is divided into domain-specific local region rather than taken holistically. Local binary pattern (LBP) is applied for the appearance-based feature description and geometric-based feature extraction. Normalized Central Moments (NCM) is used. subsections. Finally, the concatenation of LBP and NCM features is done to classify facial expression using SVM. Ryu et al. (2017) proposed a new feature called local directional ternary pattern (LDTP) to recognize facial expression. Some Edge-based methods have weakness in the smooth region. To eradicate this weakness and for the robustness in the edge region to detect edge pattern, LDTP includes encoded information about expression-related features (i.e., eyebrows, eyes, upper nose, and mouth) in which the directional information and ternary patterns are utilized. This work also coined a two-stage grid for the creation of face descriptor which includes stable codes and active codes for non-expression and expression features respectively. Due to which, this multi-stage method describes facial expression efficiently. person dependent and independent cross-validation schemes are used for the evaluation of the performance of this work.

In the past few years, various research has been performed in the area of machine learning in which deep learning has become popular. Zhang et al. (2009) introduced an approach of applying 3DCNN to manage changes in facial expressions because of muscle variation. Jung et al. (2015b) gave a deep learning method for the extraction of temporal geometry and temporal appearance of the features. Also, in their other work Jung et al. (2015a), a two-step approach is introduced. Initially, for face detection, haar-like features are applied and finally two deep networks: Convolutional Neural Network (CNN) and Deep Neural Network (DNN) is applied and this work experimentally proved that the recognition rate of CNN is better than DNN. Spiers (2016) implemented pure CNN and produced better results. An identity-aware CNN (IACNN) is
introduced in Meng et al. (2007) and two-stream IACNN is implemented in this work as separate streams. To compare different expressions, An AU-inspired deep network (AUDN) is proposed by Liu et al. (2014) in which an expression-dependent contrastive loss function is determined and in this work facial action units is used to represent facial expressions. Liu et al. (2015) utilized logistic regression and a deep belief network in Facial Expression Analysis which is based on face parsing. In this work, the face is detected then eyes, nose, and mouth extracted for facial expression study. A CNN based method is introduced in Lv et al. (2014) to identify facial expressions from candid images. These techniques lack in the training dataset and usually works on big data which are some shortcomings of these methods. Therefore, due to the small datasets of facial expression, recognition becomes challenging in the deep learning process.

3. PROPOSED FRAMEWORK

The proposed approach consists of two phases: the training phase and the testing phase. The training phase again contains five stages. In the first stage, the face region is cropped from a given image and extracted the facial components. Section 3.1 gives the detail of this stage. In the second stage, feature extraction is performed using a texture descriptor. The extracted texture patterns are then analysed using a local descriptor which results in the reduction of the dimensionality of the features as well as higher expression recognition performance as described in section 3.2 and section 3.3. In the third stage, both the appearance features are fused as described in section 3.4. Finally, for classifying the facial expression, a Multiclass Support Vector Machine (MSVM) is utilized, which is discussed in section 3.5. In the testing phase, a query image is taken as input, and all the steps mentioned in the training phase have been carried out for classifying facial expression. The proposed FER system framework is shown in Fig. 2.

An algorithm for the same is also given below:

Figure 2. The framework of the proposed FER
Algorithm:

1. **Input:** Facial Images $I(I_1, I_2, I_3, \ldots, I_n)$ representing facial expressions
2. $I_r = \text{Face detection using Viola & Jones Algorithm from } I$.
3. $F_{Ir} = \text{Extraction of facial features \{Left Eye, Right Eye, Forehead, Nose, Mouth, Right Cheek, Left Cheek\} using Eye Center Detection Algorithm.}$
4. Compute $\text{HOG}_i (F_{Ir})$.
5. $[f_1, f_2, f_3, \ldots, f_m] = \text{HOG}_i (F_{Ir})$ where $m$= Number of HOG features
6. Compute $\text{LDOP}_i (F_{Ir})$.
7. $[g_1, g_2, g_3, \ldots, g_n] = \text{LDOP}_i (F_{Ir})$ where $n$= Number of LDOP features
8. Concatenate $\text{HOG}_i (F_{Ir})$ Features & $\text{LDOP}_i (F_{Ir})$ features and store it.
9. Repeat steps from 2-8 for N input images.
10. Train classifiers for all expressions
11. Perform Classification

### 3.1 Face Detection and Regions of Interest Extraction

This is the initial stage of the proposed approach. This stage includes two steps:

1. Face Detection
2. Facial Components Extraction

#### 3.1.1 Face Detection

Almost all facial expression recognition systems include face detection as a significant step and play a vital role in recognition. The famous Voila & Jones algorithm (Voila & Jones, 2004) is being used to crop faces from the query image.

#### 3.1.2 Facial Components Extraction

The facial components include the nose, eyes, and mouth, etc. These components play an important role in detecting any facial expression. To extract facial components a method given by Viola & Jones (2001) is used and the eye centers are located in the faces using the eye centers detection algorithm (Sharma et al., 2019). Distance between the centers of the eyes is evaluated to handle head deflections also and then seven facial components viz. forehead, mouth, nose, both the eyes and cheeks are cropped.

### 3.2 Texture Feature Extraction

A Histogram of Oriented Gradients(HOG) is computed for the feature extraction process.

Mainly to represent the structural shape and appearance of an object HOG descriptors are used in an image, making them excellent descriptors for object classification. To make good texture descriptor HOG catches local intensity gradients and edge directions. HOG, developed by (Sharma et al., 2019), is a famous descriptor for object detection like SIFT (Dalal & Triggs, 2005) and SURF (Lowe, 2004). This algorithm computes the visibility of a pixel and the edge’s direction. Being a robust descriptor, HOGs represents each pixel in the region of interest of the given image to extract texture features. The imperative idea of the HOG descriptor is that directional change in the intensity represents the shape and local object appearance of an image. In HOG, given image is divided into small same-sized pieces called cells. These cells are made up of pixels. For each pixel, x and y derivatives are calculated to compute the gradient orientations. Then for each pixel, angle and gradient magnitude are being computed, and nine bins histogram quantization is calculated to form a feature vector within a cell. Similarly, the histograms will be computed for each cell within a block. Finally, resultant histograms from all blocks within an image are consolidated to form an overall feature vector.
3.3 Pattern Analysis Using a Local Descriptor

The local descriptors have been used widely because of their higher discriminative capabilities. Also, the performance of the descriptor improved due to the consideration of the multi-scale local neighbourhood; however, dimensionality cost increased. The local directional order is computed using intensity indexes at various scales in a distinctive direction. In this way, this work takes advantage of the multi-scale local neighbourhood except increasing the dimensionality of the descriptor. Originally, Dubey & Mukherjee, (2020) introduced a Local Directional Order Pattern (LDOP) descriptor for face retrieval, which outperformed previous works. A multi-radius relationship factor among the neighbouring pixels in a distinctive direction is calculated and termed as Local Directional Order. LDOP will be calculated for every pixel by obtaining the association among local directional order indexes and the center pixel. Then, the center pixel’s value will be transformed within the range of directional orders. Lastly, to build the descriptor, LDOP histogram is computed. Previously this work applied holistically over an image for the face retrieval, but in the proposed work, LDOP is being applied over facial landmarks such as the forehead, both eyes, nose, both cheeks, mouth, etc. and all the features are concatenated to form a feature vector.

3.4 The Fusion of Appearance Features

The performance of the recognition process may be influenced by the varying feature points corresponding to facial landmarks. Appearance features are interrelated, so we need to analyze local information. Thus, in the proposed approach, we have consolidated the information of both of the features of LDOP and HOG. LDOP maps provide better information regarding edge description and the appearance information to increase the discriminative abilities of the descriptor. In the proposed approach, extracted feature points from the steps of Histogram of Gradients method and Local direction order pattern are fused to form a feature vector. Since LDOP is slightly flexible to noise as it encodes the wider neighbourhood at multi-scales. Therefore, LDOP is being fused into a single feature vector with HOG to make the proposed approach robust to noise and also reduce the misclassification rate in facial expression recognition. These steps are illustrated in Fig. 3.

3.5 Classification

Classification will be the final stage in the facial expression analysis. Many supervised learning techniques have been used for classification among those techniques, Support Vector Machine (SVM) (Cortes & Vapnik, 1995) is experimentally proven as most successful for multiclassification such as facial expressions in the presence of several restrictions. That’s why SVM classifier is used in the

**Figure 3. Fusion of Appearance features**

![Diagram showing the steps of feature extraction and fusion](image-url)
proposed work for classification of the facial expressions. SVM creates a hyperplane partitioning training data into two different classes while mapping into a higher dimensional space.

Considering the problem having more than two classes, a multiclass SVM is used in the proposed work for multi-classification of facial expressions. Multi-classification classifies features into more than two classes rather than originally composed SVM for binary classification. There are two methods which are used usually for reducing multi-classification problems into multiple binary classification problems, i.e., one to one and one to all. In the proposed method, the one to all approach is used. There are three kernel functions which are extensively used in SVM: polynomial, linear and Radial Basis Function (RBF). RBF kernel function is employed in the proposed work. The parameters are picked in the proposed approach based on the work present in Yang et al. (2018). For the evaluation, 10-fold cross-validation is utilized, and the values providing the best outcome is picked. There are two stages in SVM classification: Training and testing. The proposed work classifies facial expressions into the following five classes of sentiments (shown in Fig. 4): 1: Happy 2: Sad 3: Angry 4: Disgust 5: Surprise

In the training process, Sample images are used for the construction of the feature vector which used for training and SVM mapped these feature vector in a manner in which these feature matrices refer to distinct space. When a query image is tested for the recognition of facial expression, the feature matrix of the query image is estimated, and mapping is performed on a similar space constructed in training.

4. RESULT AND DISCUSSIONS

The evaluation of the proposed facial expression recognition system is performed on several well-known datasets. Additionally, a comprehensive comparative analysis is presented with previous FER methods, conducting both qualitative and quantitative evaluations. Four datasets are employed, and each dataset in the proposed method is randomly distributed into training and testing sets.

4.1 Datasets and Configuration

The proposed facial expression recognition system is implemented using MATLAB R2015a at AMD A4-3330MX APU with Radeon(tm) HD Graphics 2.20 GHz with 4.00 GB of RAM. The included datasets are explained as follows.

Figure 4. Classification of datasets
4.1.1 Extended Cohn-Kanade Database(CK+) (Yang et al. 2017)

CK database (Jung et al. 2015a) was the first version of CK+ database having 486 images from 97 subject sequences which are FACS coded images. CK+ database is composed of both non-posed and posed expressions images from 123 subjects. CK+ dataset is updated using 107 image frames from 26 subjects. Now, CK+ dataset has 593 image sequences from 123 subjects which are captured with changing time duration. To evaluate the proposed work experimentally, 510 images are taken into account. Some sample image sequences are shown in Fig. 5(a).

4.1.2. Japanese Female Facial Expression Database (JAFFE) (Lucey et al. 2010)

There are 213 facial expression images of 10 Japanese female models in the dataset and each female model seven register their basic facial expressions. There are six expressions and one non-expression of each female model are recorded in this database. Three or four images are recorded for each model for each expression out of seven. For experimental evaluation, we used 160 images. Some example of images is shown in Fig. 5(b).

4.1.3. MMI (Lyons et al. 1998)

This database consists of more than 20 subjects including both students and research staff members having both genders out of which 44% are females. The age group of the included members is 16 to 62 years. Both males and females were from different continents (from Europe, Asia, South America, etc.) and they learnt basic expressions were taught to them and several frames having various expressions
are extracted from videos. For experimental evaluation of the proposed work, 424 images related to 5 expression categories are used. Some example of images is shown in Fig. 5(c).

### 4.1.4 Created Database

We have created our dataset to validate the proposed approach. The created dataset contains 5 classes of facial expressions viz., Happy, Surprised, Angry, Sad, Disgust, having numerous variations the age groups of people and in illumination conditions as well. The dataset is partitioned into two segments: Training Dataset and Testing Dataset. There are 500 images each in both the partitions having various facial expressions. Some sample images from this dataset is illustrated in Fig. 7. For constructing this dataset, our family members, friends, students, college workers, professors, and college guards having different age group and both of the genders (male and female) are asked to express their emotions.

### 4.2 Stage-1 Outcomes

The result analysis of stage-1, i.e. face detection and facial components extraction, is carried out in this section. Face detection and regions of interest extraction are done one after another, which are explained in section 3.1. Accordingly, the outcomes of these steps are shown in Fig. 6.

#### Figure 6. Outcomes of Stage 1: Face detection and facial components extraction

#### Figure 7. Comparison of the proposed approach with state-of-the-art methods
4.3 Stage-2 Outcomes

After the extraction of the facial components, the HOG features and LDOP features are calculated for all the facial components viz. forehead, mouth, left eye, right eye, nose, left cheek and right cheek. Then these features are concatenated to form a single feature vector.

4.4 Result Analysis of Stage 5

The Proposed work uses an integration of two appearance-based features i.e. HOG and LDOP. To assess the capability of the proposed work, some metrics are included for evaluation: True positive (TP), False Positive (FP), False Negative (FN), True Negative (TN), Then based on basic metrics the dependent metrics such as Precision, Recall, Accuracy, Misclassification Rate (MR) and F-score and has been assessed. These metrics are calculated as:

\[ Precision = \frac{TP}{TP + FP} \]  
\[ Recall = \frac{TP}{TP + FN} \]  
\[ f\text{-score} = 2 \times \frac{[Precision \times Recall]}{Precision + Recall} \]  
\[ Accuracy = \left[ \frac{TP + TN}{TP + FP + TN + FN} \right] \times 100 \]  
\[ Misclassification\ Rate = 1 - Recall \]

In Table 1, the proposed work is evaluated quantitatively with the values of different matrices on various datasets such as CK+, JAFFE, MMI and CD. It is observed that applying appearance features i.e. HOG and LDOP, which are discussed in Section 3.2 and 3.3, individually increased the number of misclassified expressions. Hence, these features are evaluated individually and images are classified. Therefore, integration of these features i.e. HOG and LDOP is done to create a single feature vector for the classification process for minimizing the misclassification rate of expression images.

Table 1. Confusion Matrix for CK+

|        | Happy | Sad | Angry | Disgust | Surprise |
|--------|-------|-----|-------|---------|----------|
| Happy  | 82    | 2   | 1     | 2       | 1        |
| Sad    | 2     | 83  | 0     | 0       | 2        |
| Angry  | 0     | 1   | 82    | 0       | 0        |
| Disgust| 2     | 0   | 2     | 68      | 1        |
| Surprise| 3    | 0   | 0     | 0       | 74       |
4.5 Comparative Evaluation of Proposed Work With State-of-the-Art Approaches

Extended Cohn–Kanade database (CK+) has been used to compare the average recognition accuracy with existing methods and the robustness of the existing techniques is assessed as illustrated in Table 1. Table 1 shows a confusion matrix for CK+ dataset. Table 2 illustrated the evaluation of the proposed approach over standard datasets. Accuracy of the proposed approach for CK+, JAFFE, MMI and Created datasets are 97.5%, 88%, 89% and 94% respectively. Furthermore, recognition rates of JAFFE, MMI and created datasets are less as compared to CK+. Therefore, the misclassification rate is high in other datasets than CK+. For Comparative evaluation of the proposed approach with existing methods, recognition rates of existing methods are compared with the proposed work and it is being observed that accuracy of the proposed approach is improved as compared to existing approaches in the field of FER. Table 3 shows the comparison of the proposed approach with the existing approaches and Fig. 7 illustrates that the proposed approach has higher accuracy on CK+ database.

5. CONCLUSION

In this paper, a fusion approach is proposed for recognizing Facial Expression. The proposed solution combines appearance features from face regions locally. This expression recognition system gives a better result than the traditional methods, i.e. holistic facial expression representation. To achieve significant performance improvement and for reduced dimensionality of the feature descriptor, a fusion of Histogram of Gradients with Local Directional Order Pattern is introduced. Local face regions have a significant amount of discriminating information which helps in the classification of facial expression. Several experiments are performed on the CK+, JAFFE, MMI and CD datasets to validate the utility of the proposed FER system for Five expressions viz. Happy, Surprise, Sad, Angry, Disgust. The quantitative result analysis showed that this method outperforms the state-of-the-art method and gives better results as compared to them. In future, this fusion descriptor can be used to analyze facial expressions having a moustache/mole/scar mark on the face. It can be extended to detect facial expressions for such images.

Table 2. Evaluation of the proposed approach over different datasets

| S.No. | Datasets | No.of images | TP  | TN  | FP  | FN  | Precision | Recall | f-score | Accuracy | MR    |
|-------|----------|--------------|-----|-----|-----|-----|-----------|--------|---------|----------|-------|
| 1     | CK+      | 410          | 389 | 11  | 6   | 4   | 0.985     | 0.988  | 0.986   | 0.975    | 0.012 |
| 2     | JAFFE    | 160          | 136 | 6   | 11  | 7   | 0.921     | 0.952  | 0.936   | 0.88     | 0.048 |
| 3     | MMI      | 359          | 311 | 8   | 16  | 24  | 0.951     | 0.928  | 0.939   | 0.89     | 0.072 |
| 4     | CD       | 2000         | 1852| 34  | 77  | 37  | 0.959     | 0.979  | 0.968   | 0.94     | 0.021 |

Table 3. Comparison of the proposed approach with state of art methods on CK+ database

| References        | Methodology               | Accuracy (%) |
|-------------------|---------------------------|--------------|
| Liu et al. 2017   | LBP +HOG                  | 96.60        |
| Shen et al. 2018  | Double δ-LBP              | 95.50        |
| Ryu et al. 2017   | LDTP                      | 94.00        |
| Yang et al. 2018  | LBP + geometric features  | 95.67        |
| Sharma et al. 2019| SIFT+MLBP                 | 97.00        |
| Proposed          | HOG+LDOP                  | 97.50        |
REFERENCES

Arshid, S., Hussain, A., Munir, A., Nawaz, A., & Aziz, S. (2018). Multi-stage binary patterns for facial expression recognition in real world. Cluster Computing, 21(1), 323–331. doi:10.1007/s10586-017-0832-5

Bartlett, M. S., Movellan, J. R., & Sejnowski, T. J. (2002). Face recognition by independent component analysis. IEEE Transactions on Neural Networks, 13(6), 1450–1464. doi:10.1109/TNN.2002.804287 PMID:18244540

Bartlett, M. S., Littlewort, G., Frank, M., Lainscsek, C., Fasel, I., & Movellan, J. (2005, June). Recognizing facial expression: machine learning and application to spontaneous behavior. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05) (Vol. 2, pp. 568–573). IEEE. doi:10.1109/CVPR.2005.297

Bay, H., Tuytelaars, T., & Van Gool, L. (2006, May). Surf: Speeded up robust features. In European conference on computer vision (pp. 404-417). Springer.

Chanti, D. A., & Caplier, A. (2018). Spontaneous facial expression recognition using sparse representation. arXiv preprint arXiv:1810.00362.

Corneanu, C. A., Simón, M. O., Cohn, J. F., & Guerrero, S. E. (2016). Survey on rgb, 3d, thermal, and multimodal approaches for facial expression recognition: History, trends, and affect-related applications. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(8), 1548–1568. doi:10.1109/TPAMI.2016.2515606 PMID:26761193

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297. doi:10.1007/BF00994018

Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR’05) (Vol. 1, pp. 886-893). IEEE.

Dols, J. M. F., & Russell, J. A. (Eds.). (2017). The science of facial expression. Oxford University Press. doi:10.1093/acprof:oso/9780190613501.003.0024

Dubey, S. R., & Mukherjee, S. (2020). LDOP: Local directional order pattern for robust face retrieval. Multimedia Tools and Applications, 79(9), 6363–6382. doi:10.1007/s11042-019-08370-x

Ekman, P., & Friesen, W. V. (1978). Facial action coding system. Investigator’s guide. Consulting Psychologists Press.

Faraji, M. R., & Qi, X. (2016). Face recognition under varying illuminations using logarithmic fractal dimension-based complete eight local directional patterns. Neurocomputing, 199, 16–30. doi:10.1016/j.neucom.2016.01.094

Ghimire, D., Jeong, S., Lee, J., & Park, S. H. (2017). Facial expression recognition based on local region specific features and support vector machines. Multimedia Tools and Applications, 76(6), 7803–7821. doi:10.1007/s11042-016-3418-y

Guo, G., & Dyer, C. R. (2003, June). Simultaneous feature selection and classifier training via linear programming: A case study for face expression recognition. In 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Proceedings., 1, I–I.

Guo, M., Hou, X., Ma, Y., & Wu, X. (2017). Facial expression recognition using ELBP based on covariance matrix transform in KLT. Multimedia Tools and Applications, 76(2), 2995–3010. doi:10.1007/s11042-016-3282-9

Jung, H., Lee, S., Yim, J., Park, S., & Kim, J. (2015a). Joint fine-tuning in deep neural networks for facial expression recognition. In Proceedings of the IEEE international conference on computer vision (pp. 2983-2991). doi:10.1109/ICCV.2015.341

Jung, H., Lee, S., Park, S., Kim, B., Kim, J., Lee, I., & Ahn, C. (2015b). Development of deep learning-based facial expression recognition system. In 2015 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV) (pp. 1–4). IEEE. doi:10.1109/FCV.2015.7103729

Liang, J., Hou, Z., Chen, C., & Xu, X. (2016). Supervised bilateral two-dimensional locality preserving projection algorithm based on Gabor wavelet. Signal, Image and Video Processing, 10(8), 1441–1448. doi:10.1007/s11760-016-0950-1
Liu, M., Li, S., Shan, S., Wang, R., & Chen, X. (2014, November). Deeply learning deformable facial action parts model for dynamic expression analysis. In Asian conference on computer vision (pp. 143-157). Springer.

Liu, M., Li, S., Shan, S., & Chen, X. (2015). Au-inspired deep networks for facial expression feature learning. Neurocomputing, 159, 126–136. doi:10.1016/j.neucom.2015.02.011

Liu, Y., Li, Y., Ma, X., & Song, R. (2017). Facial expression recognition with fusion features extracted from salient facial areas. Sensors (Basel), 17(4), 712. doi:10.3390/s17040712 PMID:28353671

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2), 91–110. doi:10.1023/B:VISI.0000029664.99615.94

Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010, June). The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In 2010 IEEE computer society conference on computer vision and pattern recognition-workshops (pp. 94-101). IEEE.

Lv, Y., Feng, Z., & Xu, C. (2014, November). Facial expression recognition via deep learning. In 2014 international conference on smart computing (pp. 303-308). IEEE.

Lyons, M. J., Akamatsu, S., Kamachi, M., Gyoba, J., & Budynek, J. (1998). The Japanese female facial expression(JAFFE) database. Proceedings of third international conference on automatic face and gesture recognition, 14-16.

Meng, Z., Liu, P., Cai, J., Han, S., & Tong, Y. (2017, May). Identity-aware convolutional neural network for facial expression recognition. In 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017) (pp. 558-565). IEEE. doi:10.1109/FG.2017.140

Nazir, M., Jan, Z., & Sajjad, M. (2018). Facial expression recognition using histogram of oriented gradients based transformed features. Cluster Computing, 21(1), 539–548. doi:10.1007/s10586-017-0921-5

Nigam, S., Singh, R., & Misra, A. K. (2018). Efficient facial expression recognition using histogram of oriented gradients in wavelet domain. Multimedia Tools and Applications, 77(21), 28725–28747. doi:10.1007/s11042-018-6040-3

Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. Information Fusion, 37, 98–125. doi:10.1016/j.infus.2017.02.003

Revina, I. M., & Emmanuel, W. S. (2018). A survey on human face expression recognition techniques. Journal of King Saud University-Computer and Information Sciences.

Ryu, B., Rivera, A. R., Kim, J., & Chae, O. (2017). Local directional ternary pattern for facial expression recognition. IEEE Transactions on Image Processing, 26(12), 6006–6018. doi:10.1109/TIP.2017.2726010 PMID:28708558

Sharma, M., Jalal, A. S., & Khan, A. (2019). Emotion recognition using facial expression by fusing key points descriptor and texture features. Multimedia Tools and Applications, 78(12), 16195–16219. doi:10.1007/s11042-018-7030-1

Shen, F., Liu, J., & Wu, P. (2018, April). Double δ-LBP: A Novel Feature Extraction Method for Facial Expression Recognition. In Chinese Conference on Image and Graphics Technologies (pp. 370-379). Springer. doi:10.1007/978-981-13-1702-6_37

Spiers, D. L. (2016). Facial emotion detection using deep learning (Doctoral Dissertation). UPPSALA Universitet.

Subramanyam, P. S., & Fegade, S. A. (2013). Face and Facial Expression Recognition-A Comparative Study. International Journal of Computer Science and Mobile Computing, 2(1).

Tian, Y., Kanade, T., & Cohn, J. F. (2011). Facial expression recognition. In Handbook of face recognition (pp. 487–519). Springer. doi:10.1007/978-0-85729-932-1_19

Valstar, M. F., Patras, I., & Pantic, M. (2005, September). Facial action unit detection using probabilistic actively learned support vector machines on tracked facial point data. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)-Workshops (pp. 76-76). IEEE. doi:10.1109/CVPR.2005.457
Valstar, M., & Pantic, M. (2006, June). Fully automatic facial action unit detection and temporal analysis. In 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW’06) (pp. 149-149). IEEE. doi:10.1109/CVPRW.2006.85

Valstar, M., & Pantic, M. (2010, May). Induced disgust, happiness and surprise: an addition to the mmi facial expression database. In Proc. 3rd Intern. Workshop on EMOTION (satellite of LREC): Corpora for Research on Emotion and Affect (p. 65). Academic Press.

Viola, P., & Jones, M. (2001, December). Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001 (Vol. 1, pp. I-I). IEEE. doi:10.1109/CVPR.2001.990517

Viola, P., & Jones, M. J. (2004). Robust real-time face detection. International Journal of Computer Vision, 57(2), 137–154. doi:10.1023/B:VISI.0000013087.49260.fb

Yang, B., Cao, J. M., Jiang, D. P., & Lv, J. D. (2017). Facial expression recognition based on dual-feature fusion and improved random forest classifier. Multimedia Tools and Applications, 77(16), 20477–20499. doi:10.1007/s11042-017-5489-9

Yang, J., Li, M., Zhang, L., Han, S., Wang, X., & Wang, J. (2018, August). Face Expression Recognition Using Gabor Features and a Novel Weber Local Descriptor. In Chinese Conference on Biometric Recognition (pp. 265-274). Springer. doi:10.1007/978-3-319-97909-0_29

Zhang, B., Gao, Y., Zhao, S., & Liu, J. (2009). Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor. IEEE Transactions on Image Processing, 19(2), 533–544. doi:10.1109/TIP.2009.2035882 PMID:19887313

Zhang, Z. (1999). Feature-based facial expression recognition: Sensitivity analysis and experiments with a multilayer perceptron. International Journal of Pattern Recognition and Artificial Intelligence, 13(06), 893–911. doi:10.1142/S0218001499000495

Zhao, X., & Zhang, S. (2016). A review on facial expression recognition: Feature extraction and classification. IETE Technical Review, 33(5), 505–517. doi:10.1080/02564602.2015.1117403