INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

Check for updates

OPEN ACCESS

Received: 14/10/2026

Accepted: 30/10/2026

Published: 11/11/2026

Citation: Poornima CK, Rashmi BS, Nandini HM (2026) Facial Expression Recognition Using Intensity-Adaptive Local Binary Pattern and Hybrid Feature Fusion. Indian Journal of Science and Technology 19(40): 2757-2767. https://doi.org/10.17485/IJST/v19i40.1362

* **Corresponding author**.

poornimateach91@gmail.com

Funding: None

Competing Interests: None

Copyright: © 2026 Poornima et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Facial Expression Recognition Using Intensity-Adaptive Local Binary Pattern and Hybrid Feature Fusion

C K Poornima[1], B S Rashmi[1]*, H M Nandini[1]

1 Department of Studies and Research in Information Technology, Karnataka State Open University, Mysuru, India

Abstract

Objectives: The main objective is to explore and develop an Intensity-Adaptive Local Binary Pattern (IA-LBP) handcrafted texture descriptor along with Sobel and Histogram of Oriented Gradient (HOG) feature fusion. This approach aims to enhance recognition of seven facial emotion classes such as anger, disgust, fear, happiness, sadness, surprise and neutral across diverse FER benchmark datasets. **Method:** The proposed IA-LBP handcrafted texture descriptor is integrated with Sobel and HOG via feature fusion analogy. IA-LBP determines adaptive threshold by using neighbourhood mean intensity and Shannon entropy. Its 256-bin histogram is integrated with six statistical Sobel features and 900 HOG features resulting in 1,162-dimensional representation from facial image. The fused feature vector is standardized using Z-Score normalization. Subsequently, SMOTE strategy is employed to address class imbalance. Further emotion prediction on facial images is carried out using various classifiers namely, Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT) and K-Nearest Neighbour (KNN). Performance evaluation is done using stratified five-fold cross-validation on five FER benchmark datasets. The efficacy of the model is evaluated using standard classification metrics like accuracy, precision, recall and F1-score. **Findings:** In the proposed IA-LBP with Sobel and HOG framework, SVM recorded the highest accuracy on CK+ (99.77%), JAFFE (98.48%), FER-2013 (82.19%) and RAF-DB (91.46%), while KNN recorded the highest accuracy on KDEF (92.00%). The effectiveness of the proposed framework is demonstrated across controlled and unconstrained facial expression datasets. The results were reported with accuracy gains of 0.17, 6.68, 1.80, 5.99 and 41.96 percentage points over previous LBP-based methods on the CK+, JAFFE, KDEF, FER-2013 and RAF-DB FER bench mark datasets respectively. However, these gains should be interpreted cautiously, taking into account differences in experimental protocols used in the compared studies. **Novelty:** The study evaluates a hybrid model by fusing IA-LBP with Sobel and HOG features. This technique substantially improves facial emotion recognition by employing the adaptive texture encoding in combination with complementary gradient features.

Keywords: Facial expression recognition; IA-LBP; Sobel features; Histogram of Oriented Gradients; Machine learning

1 Introduction

Facial expressions are an important part of non-verbal communication and provide valuable information for understanding emotional states. Facial Expression Recognition (FER) has become a vital area of research in computer vision and human-computer interaction. It is extensively applied in healthcare, education and intelligent surveillance systems. Recent reviews have highlighted the challenges to achieve a reliable FER model due to the distinct differences in facial expressions. There are a lot of variations in illumination, pose, facial appearance in facial images [1], [2]. These issues have motivated the development of distinct feature extraction and classification methods to address facial emotional expressions.

LBP is a well-known handcrafted descriptor applied in facial image analysis due to its computational simplicity and capability of representing local texture patterns. Ojala et al. [3] established the framework for LBP, which is used for texture classification. Zhao and Pietikäinen [4] pursued the LBP-based dynamic texture recognition with the focus on facial expressions. Shan et al. [5] carried out a comprehensive study of LBP for facial expression recognition, proving the relevance of using local texture information for the task. Furthermore, a number of other researchers attempted to find other descriptors and various combinations of features for facial representation. Thus, Hema and Kannan [6] put forward the idea of utilizing Patch-SIFT with the ensemble approach for learning facial emotions, whereas Harakannanavar et al. [7] were focused on the study of fusion-based algorithms for facial expression recognition. Jumani et al. [8] explored the use of LBP-based features for facial emotion identification. Yang et al. [9] explored the utilization of both multi-scale LBP features and CNN-based features, whereas Deepan et al. [10] attempted to design a hybrid Gabor filter–CNN model for facial emotion recognition purposes. Altogether, the above-mentioned articles contribute to the understanding of the importance of using handcrafted descriptors and their combinations in the FER task.

Deep-learning strategies have also significantly benefited FER from learning feature representations directly from facial images. Li and Deng [1] review deep facial recognition methods, while Canal et al. [2] present a broader review on FER approaches and challenges. Mukhopadhyay et al. [11] investigate a deep learning-based FER approach based on textural features. More recently Mukhopadhyay et al. [12] have explored multi-scale LBP image fusion using Bernoulli and entropy related information for FER. Song and Liu [13] proposed a FER network with hybrid feature extraction, while Yu [14] investigated multi-granularity feature fusion with attention mechanism. Li et al. [15] have explored hierarchical global-local feature fusion through the HGLFFNet framework, and Ahmed et al. [16] have examined a vision-transformer-based approach incorporating a hybrid local-attention mechanism in the field of FER. The abovementioned studies demonstrate that the complementary nature of feature representations and utilization of feature fusion strategies are becoming crucial aspects in modern FER studies.

Recent work has investigated the adoption of texture descriptors and feature fusion work in FER. The effectiveness of the model depends on the design, complementary attributes and the evaluation protocol employed. In spite of the advances made in handcrafted features and deep learning approaches of FER, three aspects motivate the present work. First, the conventional LBP has not emerged with mean intensity with Shannon entropy for determining the threshold value. Second, to examine the feature fusion of intensity adaptive LBP descriptors with Sobel and HOG features to capture the complementary characteristics of facial images. Third, to analyse the fused representation across FER benchmark datasets under varying conditions. These aspects motivate the proposed framework as previous studies have limited attention on this.

To address this gap, an IA-LBP descriptor is proposed and fused with Sobel and HOG features to carry out FER. To devise this IA-LBP descriptor, centre pixel intensity and Shannon entropy is used for threshold determination. Based on this threshold, the eight neighbourhood pixels were encoded with binary bits and aided in generating a histogram. Using the Sobel operator, the six statistical features are extracted from horizontal, vertical and gradient magnitude from each facial image. Likewise, the distribution of gradient orientation features is extracted from HOG.

In this research work, various machine-learning classifiers like SVM, RF, DT and KNN are used to assess the proposed framework. To carry out the experimental analysis the five benchmark datasets namely, CK+, JAFFE, KDEF, FER-2013 and RAF-DB are utilized [17], [18], [19], [20], [21]. Stratified five-fold cross-validation technique is employed during experimental evaluation. During the training phase, feature standardization and Synthetic Minority Over-sampling Technique (SMOTE) are employed to prevent data leakage. The main contributions of this study are:

- To devise IA-LBP texture descriptor in determining adaptive threshold using neighbourhood mean intensity and Shannon entropy.
- To establish hybrid feature representation by integrating IA-LBP with Sobel and HOG descriptors to express discriminative facial features.
- To conduct empirical evaluation of the proposed hybrid framework using five diverse FER benchmark datasets.
- To assess FER accuracy across four machine-learning classifiers namely, SVM, RF, DT and KNN and compare using a stratified five-fold cross-validation protocol.

2 Methodology

The distinctive categorization of varied human facial emotions is tackled by integrating three mutually supportive feature extraction techniques and consequently using machine learning-based classification for facial emotion prediction. The FER framework comprises various phases as portrayed in Figure 1.

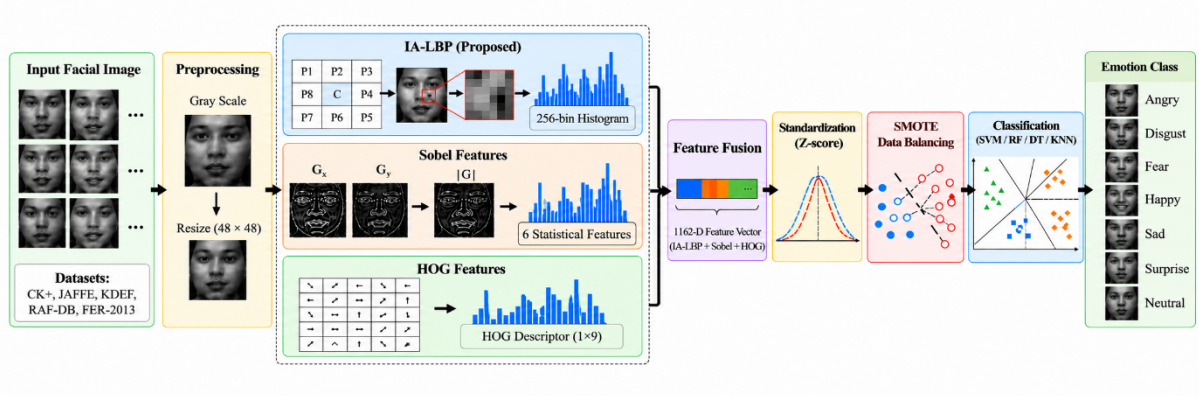


Fig 1. Overview of IA-LBP Framework for Facial Emotion Recognition.

It is a known fact that human facial images depict emotion variability such as texture patterns, geometric properties and movement of muscles. The pre-processing stage transforms facial images obtained from multiple benchmark datasets into a unified format. All the dataset images are transformed to grayscale and are downscaled to fixed resolution of 48x48 for compatible purposes.

Posterior to this, the input is passed through three parallel feature descriptors: the proposed IA-LBP, Sobel Gradient and HOG. The three complementary descriptors capture texture, edge and shape characteristics of facial expressions respectively. The integrated facial expression features are framed as a single feature vector in the fusion layer. This fused vector is propagated to machine learning classifiers where it learns the fused patterns and produces the final prediction related to various facial emotion classes.

2.1 Proposed IA-LBP Feature Extraction

Usually the human face contains smooth (cheeks, forehead) and highly textured regions (eyes, eyebrows, lips). It is tedious to describe both smooth and rough visual variability of the facial images using global texture descriptors. In this context, the LBP texture descriptor has been emphasized due to its robustness and simplicity

2.1.1. Conventional LBP

In this variant, the 3x3 neighbourhood pixels are compared with centre pixel intensity and binary encoding rule is deployed to calculate the LBP code. However, this method may not adequately capture subtle facial texture variations. However, this variant and few other LBP variants reported earlier [22], [23] employs centre pixels as fixed threshold making it sensitive to illumination fluctuations. Therefore, the proposed IA-LBP variant uses entropy, neighbourhood mean intensity and also centre pixel to compute the adaptive threshold.

2.1.2. Entropy Computation

In order to compute this, initially the normalised probability of each pixel in the facial image is computed as:

$$p_i = \frac{I_i}{\sum_{j=1}^{9} I_j} \tag{1}$$

Further, the local Shannon entropy for the P_i is then computed as:

$$H = -\sum_{i=0}^{8} p_i \log_2(p_i) \tag{2}$$

It is examined that a high value of entropy has been experienced depicting the rough region and low value for smooth region in facial images.

2.1.3. Spatial Mean Intensity

The mean intensity of i^{th} pixel in the local spatial region is computed as:

$$\mu = \frac{1}{9}\sum_{i=1}^{9} I_i \tag{3}$$

As per the local illumination criteria in the facial image, β value is formulated as the normalized intensity factor and dynamically adjust the adaptive threshold:

$$\beta = \frac{\mu}{255} \tag{4}$$

It is examined that high-luminance areas in facial image denote large β value and low-luminance denote less β value.

2.1.4. Adaptive Threshold Computation

The key significance of the proposed variant is in describing the adaptive threshold by aggregating the centre pixel intensity (C), local Shannon entropy (H) and normalized mean intensity level (β) which is expressed as:

$$T = C + \beta H \times 10 \tag{5}$$

It is examined that the constant value 10 maintains stable threshold values and reduces the influence of local texture complexity.

2.1.5. Binary Pattern Generation

Further, comparison is made with the 8-neighbouring pixel at every pixel of facial image as per the binary decision rule:

$$b_i = \begin{cases} 1, & P_i \geq T \\ 0, & P_i < T \end{cases} \tag{6}$$

The thresholded values are thus multiplied by binary weights to examine the subtle texture spots on the facial image and are encoded to generate LBPIA-LBP code which is expressed as:

$$LBP_{IA} = \sum_{i=0}^{7} b_i 2^i \tag{7}$$

The numerical illustration for IA-LBP computation is given in Figure 2 (a-e). For example, the 3×3 grayscale neighbourhood of the facial image is considered and the flow is depicted to compute the decimal IA-LBP code.

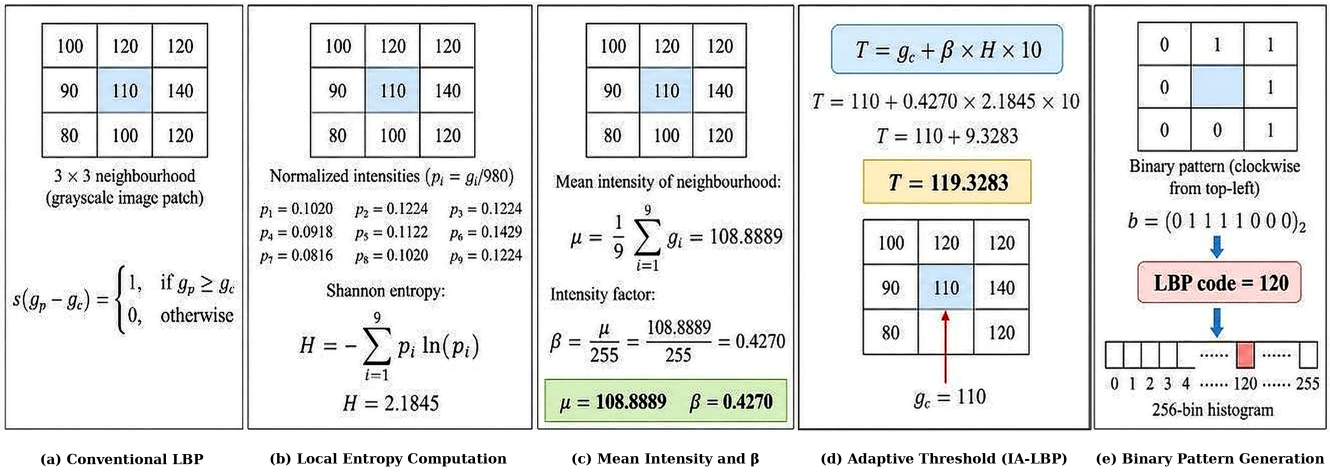


Fig 2. Numerical Illustration of proposed IA-LBP Descriptor.

2.1.6. Histogram Representation

The decimal codes generated for all pixels are accumulated into a 256-bin histogram of every facial image and normalized histogram $H_{norm}(k)$ is formulated as:

$$H_{norm}(k) = \frac{H(k)}{\sum_{i=0}^{255} H(i)} \tag{8}$$

Where *k* is histogram bin *H(i)* is counter frequency in i^{th} bin. Finally, the IA-LBP feature vector is produced using the normalized histogram values of all the consecutive facial images which is denoted as:

$$F_{IA-LBP} = \left[H_{norm(0)}, H_{norm(1)}, ..., H_{norm(N-1)}\right] \tag{9}$$

Further, this proposed IA-LBP feature vector is combined with Sobel and HOG feature vectors to construct the hybrid feature representation.

2.2 Sobel Gradient Feature Extraction

Facial expressions such as happiness, surprise and anger are indicated by substantial edge and gradient changes. Accordingly, the Sobel operator is employed to showcase facial boundaries and intensity shifts. Therefore, six statistical measures for every pre-processed grayscale facial image are computed to determine concise summary of the prominent face edge pattern as follows:

$$F_{Sobel} = [\mu_{G_x}, \sigma_{G_x}, \mu_{G_y}, \sigma_{G_y}, \mu_G, \sigma_G] \tag{10}$$

Where, the mean and standard deviation of the horizontal gradient are computed as:

$$\mu_{G_x} = \frac{1}{N}\sum_{i=1}^{N} G_x(i), \sigma_{G_x} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(G_x(i) - \mu_{G_x})^2} \tag{11}$$

Similarly, the mean and standard deviation of the vertical gradient are computed as:

$$\mu_{G_y} = \frac{1}{N}\sum_{i=1}^{N} G_y(i), \sigma_{G_y} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(G_y(i) - \mu_{G_y})^2} \tag{12}$$

Finally, the statistical properties of the gradient magnitude image are determined as:

$$\mu_G = \frac{1}{N}\sum_{i=1}^{N} G(i), \sigma_G = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(G(i) - \mu_G)^2} \tag{13}$$

Thus, the six-column feature vector for every facial image is created and thus serves as an edge component in this proposed work.

2.3 HOG Feature Extraction

At the outset, HOG[24] is emphasized in expressing structural layout of facial emotions such as anger, fear, happiness and surprise. This aids in distinguishing the local orientation and facial geometry across eye, eyebrow, lip and mouth segments.

HOG descriptor is parameterized with orientation bins as 9, cell size as 8x8, block size as 2x2 and L2-Hys as block normalization. This leads to generation of HOG feature vector for every pre-processed facial image as:

$$F_{HOG} = [h_1, h_2, ..., h_n] \tag{14}$$

This vector is fused with IA-LBP histogram vector and Sobel statistical feature vector to constitute a hybrid feature model for identifying facial expression.

2.4 Feature Fusion

To achieve the best face representation, the fusion of IA-LBP, Sobel and HOG at the feature level is carried out. Each of the pre-processed 48 × 48 grayscale facial images had a 256-bin histogram for IA-LBP feature representation. The 6 Sobel features comprised the mean and standard deviation of the horizontal gradient, vertical gradient and overall gradient magnitude. The HOG features had 900 feature vectors with 9 bins for gradient direction, 8 × 8 pixels for cell size and 2×2 cells for block size. A concatenation of the three descriptors formed a 1,162-dimensional feature vector for facial-expression classification as:

$$F_{fusion} = [F_{IA-LBP}|| F_{Sobel}|| F_{HOG}] \tag{15}$$

2.5 Feature Standardization

The fused feature vector comprises diverse features from the IA-LBP, Sobel and HOG descriptors with varying dynamic ranges and variances. Therefore, Z-score standardization is adapted to scale features reliably. During each fold of the stratified five-fold cross-validation, mean and standard deviation were computed only for the training samples. These derived parameters were then used to transform the validation samples ensuring data leakage prevention.

2.6 SMOTE-Based Data Balancing

Class imbalance in benchmark FER datasets can reduce the recognition rate of minority classes. In order to resolve this issue, SMOTE [25] is employed to generate synthetic minority class samples using interpolation strategy. SMOTE is employed exclusively in each training fold to prevent data leakage into the validation fold during the cross-validation process.

2.7 Traditional Supervised Classification

The proposed hybrid framework classification accuracy is assessed across four machine-learning classifiers namely, SVM, RF, DT and KNN by adapting stratified five-fold cross-validation. The mean classification accuracy is formulated as:

$$\bar{A} = \frac{1}{K}\sum_{i=1}^{K} A_i \tag{16}$$

where $\overline{A}$ is the accuracy of the i^{th} fold and K=5. In order to categorize the non-linear emotion features, the Radial Basis Function (RBF) kernel is employed by SVM. RF classifier used 100 decision trees using majority voting for emotion classification. In DT, the feature set is recursively partitioned for predicting the emotion classes. DT is used to classify the samples by partitioning the feature set recursively. Finally, KNN focused on nearest training neighbours using Euclidean distance for classifying the emotion samples.

2.8 Datasets

The proposed framework of FER is evaluated on five benchmark datasets representing controlled and unconstrained conditions. They differ in the diversity of subjects, image properties, acquisition conditions and facial-expression categories and are described in Table 1.

The database is divided according to the seven facial emotion classes such as anger, disgust, fear, happiness, sadness, surprise and neutral. The contempt emotion (*) class in CK+ dataset is excluded from the present study.

2.9 Experimental Environment

The proposed FER framework is implemented in the Python programming language using the Visual Studio Code software. OpenCV, NumPy, Pandas, SciPy and scikit-image have aided in mathematical and image processing operations. The fused hybrid vector is standardized using Z-score normalization and class imbalance is addressed by SMOTE. Empirical studies were performed across five benchmark datasets using SVM, RF, DT, and KNN classifiers by employing stratified five-fold cross-validation. FER classification accuracy is measured using accuracy, precision, recall, F1-score, confusion matrices, ROC curves, AUC and learning curves.

Table 1. Description of the Benchmark Dataset used for FER

Datasets	No. of subjects	Original samples	Image type	Emotion categories	Acquisition condition
CK+ Extended Cohn–Kanade Dataset [17]	123	593 sequences	Mostly grayscale	7*	Laboratory-controlled
JAFFE Japanese Female Facial Expression Database [18]	10	213 images	Grayscale	7	Laboratory-controlled
KDEF Karolinska Directed Emotional Faces [19]	70	4,900 images	Color	7	Controlled studio
FER-2013 Facial Expression Recognition 2013 [20]	-	35,887 images	Grayscale, 48×48	7	In-the-wild
RAF-DB Real-world Affective Faces Database [21]	-	29,672 images	Color	7	In-the-wild

2.10 Performance Evaluation

The performance of the proposed facial emotion classification model is evaluated using accuracy, precision, recall and F1-score across facial emotion benchmark datasets. Accuracy measures the proportion of correctly classified facial emotions.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{17}$$

Precision represents the actual correctness of the positive predictions for a specific emotion and recall measures the model's ability to recognize the actual instances of a specific emotion.

$$Precision = \frac{TP}{TP+FP}, Recall = \frac{TP}{TP+FN} \tag{18}$$

To evaluate false and missed facial emotion classification, precision and recall measures are combined to depict F1-score which is formulated as:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{19}$$

Further, confusion matrix is employed to illustrate comparison across classes. Also, Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC) is illustrated to portray the performances of the classifiers.

3 Results and Discussion

Table 2 depicts the classification results of facial emotions using four machine learning classifiers across five benchmark FER datasets with distinct levels of complexity and variability.

Table 2. Classification results of the proposed IA-LBP feature representation across benchmark FER datasets

Dataset	SVM Accuracy (%)	Random Forest Accuracy (%)	Decision Tree Accuracy (%)	KNN Accuracy (%)	Best Classifier	Precision	Recall	F1-score
CK+	**99.77 ± 0.46**	99.60 ± 0.39	85.03 ± 1.34	83.65 ± 2.29	SVM	1.00	1.00	1.00
JAFFE	**98.48 ± 0.66**	95.57 ± 0.84	75.48 ± 1.69	94.53 ± 1.53	SVM	0.98	0.98	0.98
KDEF	91.57 ± 1.32	91.57 ± 1.55	72.10 ± 3.58	**92.00 ± 1.39**	KNN	0.92	0.92	0.92
FER-2013	**82.19 ± 0.32**	78.91 ± 0.74	54.53 ± 0.28	66.72 ± 0.29	SVM	0.82	0.82	0.82
RAF-DB	**91.46 ± 0.54**	87.99 ± 0.29	61.21 ± 0.91	73.34 ± 0.56	SVM	0.92	0.91	0.91

From the Table 2 results it can be observed that SVM achieved the best accuracy for four out of five datasets namely CK+, JAFFE, FER-2013 and RAF-DB. However, the KNN recorded the best accuracy for the KDEF dataset. The highest accuracy

achieved by SVM is on CK+ with 99.77 $\pm$ 0.46% followed by JAFFE with 98.48 $\pm$ 0.66%. For CK+ dataset the SVM classifier is closely followed by the Random Forest with an accuracy of 99.60 $\pm$ 0.39%. This implies that the extracted features have excellent ability to separate emotion classes for CK+ dataset. On JAFFE, the SVM achieved an accuracy of 98.48 $\pm$ 0.66%, with precision, recall and F1-score of 0.98. This shows that the SVM classifier correctly predicted most of the examples across all facial emotion classes. It can be concluded from the above results that the extracted hybrid features possess a great discriminatory ability for the facial expression recognition of CK+ dataset.

For KDEF, KNN reaches the highest accuracy of 92.00 $\pm$ 1.39%, 0.92 precision, 0.92 recall, and 0.92 F1-score. Meanwhile, the accuracy of SVM and Random Forest classifiers is close to each other, attaining 91.57 $\pm$ 1.32% and 91.57 $\pm$ 1.55% respectively.

On FER-2013, SVM achieved an accuracy of 82.19 $\pm$ 0.32%, with a precision, recall and F1-Score of 0.82. The lower accuracy may be caused by the variations in facial appearance, poses, illumination, image quality, alignment and expression intensity in unconstrained images.

For RAF-DB, SVM yielded 91.46 $\pm$ 0.54% with a precision, recall and F1-score of 0.92, 0.91 and 0.91 respectively. The results demonstrate the effectiveness of the proposed hybrid features, while differences in performance reflect the varying complexity of the facial emotion datasets.

As can be seen from Figure 3(a-e), the validation accuracy tends to increase with the growth of the training-set size for all the five datasets, while the training accuracy maintains a high level. It can also be observed that the curves gradually converge, which implies that the learning process is stable. In addition, it can be found that larger training-validation gaps are achieved for JAFFE, FER-2013 and RAF-DB, while a better model convergence is achieved for CK+ and KDEF with sufficient training samples.

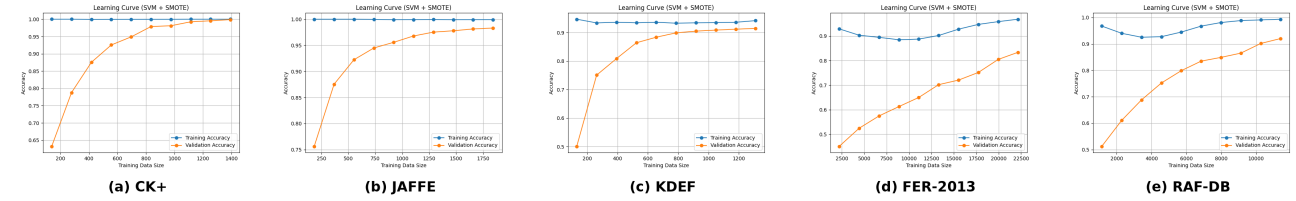


Fig 3. Learning curves of the proposed framework using SVM with SMOTE across five benchmark FER datasets.

Figure 4(a-e), displays the confusion matrices for the best classifiers on the five datasets. The majority of the diagonal elements are prominent, which implies correct classifications, particularly for CK+ and JAFFE which achieved the best accuracy scores. Similarly, the KDEF dataset is well classified with some inter-class confusions, while the FER-2013 dataset demonstrates significant inter-class confusions due to its unrestricted nature. In turn, RAF-DB has clear diagonals, which means that it does not have many misclassifications.

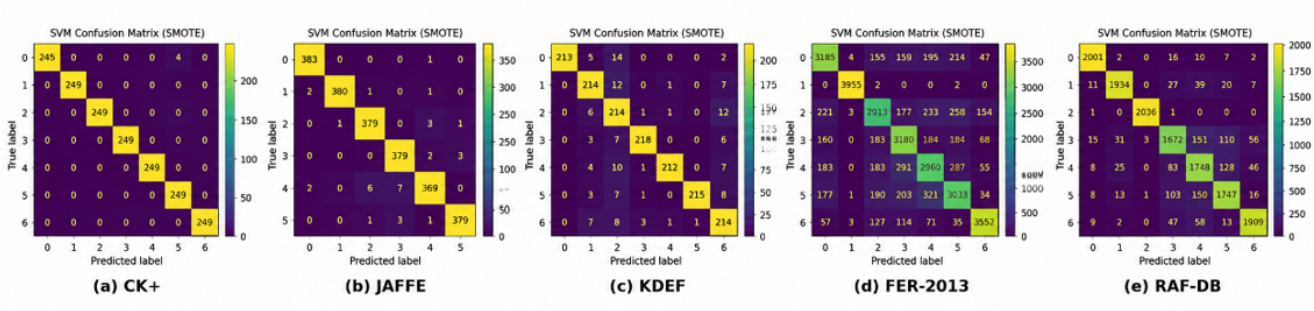


Fig 4. Confusion matrices of the proposed framework across five benchmark FER datasets.

Figure 5(a-e), shows the ROC curves of the best classifier for each benchmark dataset with AUC calculated using the predictions in the validation. All of the presented curves lie above the diagonal line indicating a good ability of the classifiers to distinguish between classes. The CK+ is close to the perfect classification, followed by the JAFFE and KDEF which demonstrate strong class discrimination, whereas RAF-DB shows comparatively lower performance. Meanwhile for FER-2013, there is a significant overlap between positive and negative classes due to a high difficulty level of this benchmark.

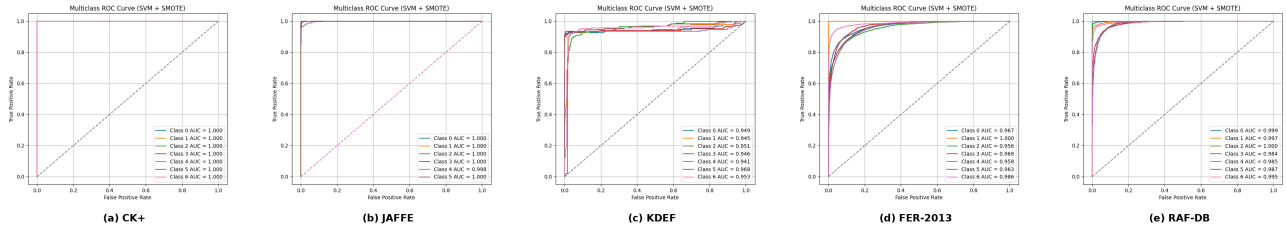


Fig 5. ROC curves of the proposed framework across five benchmark FER datasets.

3.1 Comparison of the proposed framework with recent LBP-based FER approach

Table 3 compares the proposed hybrid handcrafted feature framework with the recent LBP-based FER approach of Mukhopadhyay et al.[11],[12]. The proposed framework achieved higher numerical accuracy across all five benchmark datasets. The gains were 0.17, 6.68, 1.80, 5.99 and 41.96 percentage points for CK+, JAFFE, KDEF, FER-2013 and RAF-DB respectively. SVM obtained the best accuracy on CK+, JAFFE, FER-2013 and RAF-DB, while KNN performed best on KDEF. The results show that the proposed combination of IA-LBP, Sobel and HOG provide competitive handcrafted representations for both controlled and unconstrained FER datasets. However, the comparison must be interpreted with respect to the methodological differences. Mukhopadhyay et al.[11] employed CLBP-CNN based method, while Mukhopadhyay et al.[12] employed multi-scale LBP image fusion in conjunction with a customized CNN. In contrast, proposed framework operates on the fusion of handcrafted feature vectors followed by traditional machine-learning classifiers.

Table 3. Comparison of the proposed framework with recent LBP-based FER approach

Dataset	Mukhopadhyay et al.[11] (%)	Mukhopadhyay et al.[12] (%)	Best Proposed (%)	Improvement (pp)
CK+	91.00	99.60	**99.77 (SVM)**	+0.17
JAFFE	82.20	91.80	**98.48 (SVM)**	+6.68
KDEF	Not reported	90.20	**92.00 (KNN)**	+1.80
FER-2013	64.50	76.20	**82.19 (SVM)**	+5.99
RAF-DB	Not reported	49.50	**91.46 (SVM)**	+41.96

3.2 Contextual Comparison of the Reported FER Assessment Studies

Table 4 compares the performance of the proposed IA-LBP + Sobel + HOG framework in terms of accuracy against the five selected benchmark datasets and representative deep-learning-based facial expression recognition methods. On the CK+ dataset, the proposed framework exhibited a slightly lower accuracy than the CNN-10 model (99.95%). In turn, on the JAFFE database, the proposed method's accuracy outperformed the CNN-10 (95.41%) and VGG-19 (92.21%) approaches. On the KDEF database, the proposed KNN-based framework attained 92.00% accuracy, which is comparable to the result of DeepExp3D (92.24%) but lower than that of HSNet (93.35%) and higher than that of AFER (90.62%). On the FER-2013 database, the proposed SVM-based framework outperformed the ResMaskingNet (74.14%) and its ensemble version (76.82%) but underperformed the CNN-10 (84.30%). On the RAF-DB, the proposed method's accuracy is higher than those of DAN (89.70%) and HDCNet (90.74%). Thus, compared to the selected deep-learning methods, the handcrafted-based features attained comparable accuracy across most of the benchmark datasets. Nevertheless, the results are presented as a contextual comparison with previous FER studies rather than evidence of superiority. The reported studies employ different assessment and evaluation protocols.

IA-LBP utilizes intensity adaptive thresholds to capture local facial texture details, while Sobel and HOG extract gradient and structural cues. The fusion of the diverse cues achieved competitive results on FER tasks. However, the comparison with other approaches is made on different data splits and pre-processed data with different augmentations. Thus, the results prove the method's effectiveness but not its superiority.

Table 4. Contextual comparison of the Reported FER Assessment Studies

Dataset	Author(s), Year	Deep-learning approach	Accuracy (%)	Proposed method (%)
CK+	Dada et al. [26]	CNN-10	**99.95**	**99.77** (SVM)
		VGG-19	77.51	
JAFFE		CNN-10	**95.41**	**98.48** (SVM)
		VGG-19	92.21	
KDEF	Fan et al. [29]	HSNet	**93.35**	**92.00** (KNN)
		DeepExp3D	92.24	
		AFER	90.62	
FER-2013	Dada et al. [26]	CNN-10	**84.30**	**82.19** (SVM)
	Pham et al. [30]	ResMaskingNet	74.14	
		ResMaskingNet ensemble	76.82	
RAF-DB	Wen et al. [27]	DAN	89.70	**91.46** (SVM)
	Zhu et al. [28]	HDCNet	90.74	

4 Conclusion

This study proposed an Intensity-Adaptive Local Binary Pattern (IA-LBP) descriptor integrated with statistical Sobel and Histogram of Oriented Gradients (HOG) features for facial emotion prediction. The proposed framework combines adaptive local texture encoding with complementary edge and gradient-orientation information to construct a unified handcrafted feature representation. Its performance is evaluated using four conventional machine-learning classifiers across five benchmark datasets: CK+, JAFFE, KDEF, FER-2013 and RAF-DB using stratified five-fold cross-validation protocol.

The experimental results demonstrated that Support Vector Machine (SVM) achieved the highest classification accuracy on CK+ (99.77%), JAFFE (98.48%), FER-2013 (82.19%) and RAF-DB (91.46%), whereas KNN performed best on KDEF (92.00%). These findings indicate that the proposed hybrid framework can support facial expression classification across datasets with different characteristics and imaging conditions. Comparisons with selected LBP-based approaches showed encouraging accuracy improvements, although datasets, pre-processing and evaluation protocols differ.

The principal contribution of this work lies in combining an intensity and entropy LBP formulation with complementary Sobel and HOG descriptors and evaluating cross multiple benchmark datasets. Despite these promising results, the study has certain limitations. The evaluation is restricted to five benchmark datasets and four conventional classifiers. Hence, the handcrafted representation may not capture all the complex variations associated with facial identity, head pose, illumination, occlusion and subtle expressions in unconstrained environments. Furthermore, the computational efficiency and the individual contributions of IA-LBP, Sobel and HOG have not been fully established without dedicated ablation and runtime analyses. Future research will investigate descriptor ablation, alternative adaptive-threshold formulations and comparisons with lightweight deep-learning models. Future evaluation on a diverse set of in-the-wild datasets and demographic groups can also be carried out.

5 Acknowledgements

Nil

Author details: C. K. Poornima is a Research Scholar; B. S. Rashmi & H. M. Nandini are Assistant Professors.

References

1) Li S, Deng W. Deep facial expression recognition: a survey. *IEEE Transactions on Affective Computing*. 2022;13(3):1195–1215. https://doi.org/10.1109/TAFFC.2020.2981446.
2) Canal FZ, Müller TR, Matias JC, de Sa Junior AR, Pozzebon E, Sobieranski AC, et al. A survey on facial emotion recognition techniques: a state-of-the-art literature review. *Information Sciences*. 2022;582:593–617. https://doi.org/10.1016/j.ins.2021.10.005.
3) Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2002;24(7):971–987. https://doi.org/10.1109/TPAMI.2002.1017623.
4) Zhao G, Pietikäinen M. Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2007;29(6):915–928. https://doi.org/10.1109/TPAMI.2007.1110.
5) Shan C, Gong S, McOwan PW. Facial expression recognition based on Local Binary Patterns: a comprehensive study. *Image and Vision Computing*. 2009;27(6):803–816. https://doi.org/10.1016/j.imavis.2008.08.005.

6) Hema D, Kannan S. Patch-SIFT: enhanced feature descriptor to learn human facial emotions using an ensemble approach. *Indian Journal of Science and Technology*. 2021;14(21):1740–1747. https://doi.org/10.17485/IJST/v14i21.2261.
7) Harakannanavar SS, Sapnakumari C, Ramachandra AC, Pramodhini R, Prashanth CR. Performance evaluation of fusion based efficient algorithms for facial expression recognition. *Indian Journal of Science and Technology*. 2023;16(4):266–276. https://doi.org/10.17485/IJST/v16i4.1891.
8) Jumani SZ, Ali F, Kandhro IA, Lakhan QA, Ali U, Haroon MW, et al. Facial emotion identification based on Local Binary Pattern feature detector. *Indian Journal of Science and Technology*. 2019;12(28):1–6. https://doi.org/10.17485/ijst/2019/v12i28/146447.
9) Yang H, Gemeay E, Alqahtani A, Alanazi A, Alsubai S, Lee S. Facial expression recognition by multi-scale local binary patterns (MLBP) and convolutional neural network (CNN) features. *Expert Systems*. 2025;42(6):e70044. https://doi.org/10.1111/exsy.70044.
10) Deepan P, Vidya R, Reddy MA, Arul N, Ravichandran J, Dhiravidaselvi S. A hybrid Gabor filter-convolutional neural networks model for facial emotion recognition system. *Indian Journal of Science and Technology*. 2024;17(35):3696–3703. https://doi.org/10.17485/IJST/v17i35.1998.
11) Mukhopadhyay M, Dey A, Kahali S. A deep-learning-based facial expression recognition method using textural features. *Neural Computing and Applications*. 2023;35(9):6499–6514. https://doi.org/10.1007/s00521-022-08005-7.
12) Mukhopadhyay M, Kar C, Kahali S, Dey A. Synergizing Bernoulli and entropy for robust multi-scale LBP image fusion in facial emotion recognition. *Journal of Intelligent and Fuzzy Systems*. 2026;50(4):1215–1233. https://doi.org/10.1177/18758967251350574.
13) Song D, Liu C. A facial expression recognition network using hybrid feature extraction. *PLoS One*. 2025;20(1):e0312359. https://doi.org/10.1371/journal.pone.0312359.
14) Yu G. A multi-granularity feature fusion approach with attention for facial expression recognition. *Scientific Reports*. 2025;15:42507. https://doi.org/10.1038/s41598-025-26533-9.
15) Li H, Zhang M, Wu J, Zhang F, Sun F, Wang M. HGLFFNet: hierarchical global-local feature fusion network for facial expression recognition. *Neurocomputing*. 2026;664:132096. https://doi.org/10.1016/j.neucom.2025.132096.
16) Ahmed MA, Dong J, Shi R, Almaqtari H, et al. Vision transformer-based facial expression recognition with hybrid local attention mechanism. *Machine Vision and Applications*. 2026;37:90. https://doi.org/10.1007/s00138-026-01860-7.
17) Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z, Matthews I. 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. 2010;p. 94–101. https://doi.org/10.1109/CVPRW.2010.5543262.
18) Lyons MJ, Akamatsu S, Kamachi M, Gyoba J. Coding facial expressions with Gabor wavelets. In: Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition. 1998;p. 200–205. https://doi.org/10.1109/AFGR.1998.670949.
19) Lundqvist D, Flykt A, Öhman A. The Karolinska Directed Emotional Faces (KDEF). Stockholm. Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet. 1998.
20) Goodfellow IJ, Erhan D, Carrier PL, Courville A, Mirza M, Hamner B, et al. Challenges in representation learning: a report on three machine learning contests. In: Neural Information Processing: 20th International Conference, ICONIP 2013. Berlin. Springer. 2013;p. 117–124. https://doi.org/10.1007/978-3-642-42051-1_16.
21) Li S, Deng W, Du J. Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition. 2017;p. 2584–2593. https://doi.org/10.1109/CVPR.2017.277.
22) Rashmi BS, Nagendraswamy HS. Video shot boundary detection using midrange local binary pattern. In: Proceedings of the 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI). Jaipur, India. 2016;p. 201–206. https://doi.org/10.1109/ICACCI.2016.7732047.
23) Nandini HM, Chethan HK, Rashmi BS. Abrupt shot change detection using Midhinge Local Binary Pattern. In: 2020 IEEE Hyderabad Section International Conference (IEEE-HYDCON). 2020;p. 1–5. https://doi.org/10.1109/HYDCON48903.2020.9242841.
24) Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). 2005;p. 886–893. https://doi.org/10.1109/CVPR.2005.177.
25) Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*. 2002;16:321–357. https://doi.org/10.1613/jair.953.
26) Dada EG, Oyewola DO, Joseph SB, Emebo O, Oluwagbemi OO. Facial emotion recognition and classification using the convolutional neural network-10 (CNN-10). *Applied Computational Intelligence and Soft Computing*. 2023;2023:2457898. https://doi.org/10.1155/2023/2457898.
27) Wen Z, Lin W, Wang T, Xu G. Distract your attention: multi-head cross attention network for facial expression recognition. *Biomimetics*. 2023;8(2):199. https://doi.org/10.3390/biomimetics8020199.
28) Zhu X, Sun J, Liu G, Shen C, Dai Z, Zhao L. Hybrid domain consistency constraints-based deep neural network for facial expression recognition. *Sensors*. 2023;23(11):5201. https://doi.org/10.3390/s23115201.
29) Fan X, Jiang M, Shahid AR, Yan H. Hierarchical scale convolutional neural network for facial expression recognition. *Cognitive Neurodynamics*. 2022;16(4):847–858. https://doi.org/10.1007/s11571-021-09761-3.
30) Pham L, Vu TH, Tran TA. Facial expression recognition using residual masking network. In: Proceedings of the 25th International Conference on Pattern Recognition (ICPR). Milan, Italy. 2021;p. 4513–4519. https://doi.org/10.1109/ICPR48806.2021.9411919.