© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715104

Facial Emotion Recognition Using Local Directional Ternary Patterns and SVM

A. SRIKRISHNA[1], L. SAI SRI KUMAR[2], J. SANDHYA[3], G. KARTHIK[4], D. BINDU MADHAVI BAI[5]
[1,2,3,4,5]*Department of Information Technology, R.V.R & J.C College of Engineering, Guntur, India*

Abstract- Facial Expression Recognition (FER) is a crucial component of human-computer interaction, affective computing, and behavioural analysis systems. Although deep learning-based methods have shown impressive results, they are often data-intensive and computationally expensive. In contrast, handcrafted texture descriptors remain useful for FER tasks with limited data. In this paper, we present an improved Local Directional Ternary Pattern (LDTP)-based facial expression recognition system that aims to enhance directional texture pattern encoding and spatial feature aggregation for better discriminative power. The proposed method focuses on the major directional patterns and combines spatially localized feature description over facial regions, allowing for better expression-related subtle variations. The LDTP features are then modelled using a Support Vector Machine classifier with a radial basis function kernel. Experiments on the CK+ database with an image-level stratified train-test split show that the proposed method can achieve an accuracy of 89.34%, which is better than the baseline LBP- and traditional LDTP-based methods under the same experimental setup. These findings suggest that the proposed LDTP-based approach is an effective and efficient solution for facial expression recognition.

Index Terms- Local Directional Ternary Pattern, Feature Extraction, Emotion Recognition, Texture Descriptors

I. INTRODUCTION

Facial Expression Recognition (FER) is a basic task in computer vision and affective computing, seeking to automatically detect human emotional expressions based on the movement of facial muscles. Accurate FER systems are critical for various tasks, including human-computer interaction, behavioural analysis, intelligent surveillance, and mental state evaluation. Despite much research, FER remains a difficult task due to the differences between and within expression classes, as well as robustness issues with respect to illumination, pose, and image resolution.

Recent breakthroughs in deep learning have greatly promoted the accuracy of FER systems. Nevertheless, these methods generally demand large-scale annotated image databases, intensive computational support, and extensive hyperparameter tuning [11], [14], [15]. By contrast, hand-designed texture descriptors remain useful in applications where data are scarce and computational complexity is a concern. Among them, Local Binary Patterns (LBP) and Local Directional Patterns (LDP) are popular FER methods [1], [2], [8]. Although LBP is designed to describe local intensity patterns and LDP concentrates on edge direction features, both methods employ binary encoding schemes, which are extremely noise- and intensity-sensitive, especially for low-resolution facial images.

To overcome these drawbacks, the Local Directional Ternary Pattern (LDTP) proposes ternary thresholding of directional responses, which enables the suppression of weak directional responses and retains the major expression-related directions [4], [5]. Based on this idea, this paper proposes an improved LDTP-based facial expression recognition system that refines the selection of directional responses and combines spatially localized feature aggregation over facial regions. By focusing on the most discriminative directional patterns and ensuring efficiency, the proposed method enhances the capability of LDTP in FER. Classification is done using a Support Vector Machine with a radial basis function kernel.

II. METHODOLOGY

This paper proposes an improved face expression recognition system, integrating a step-by-step refined Local Directional Ternary Pattern feature extraction process with a Support Vector Machine classifier as depicted in Fig. 1. The overall approach includes preprocessing, baseline LDTP feature extraction, improved LDTP feature representation, formation of the feature vectors, and classification. Each of these components is designed to introduce more robustness to noise, illumination changes, and minute facial muscular activity with a lower computational complexity.

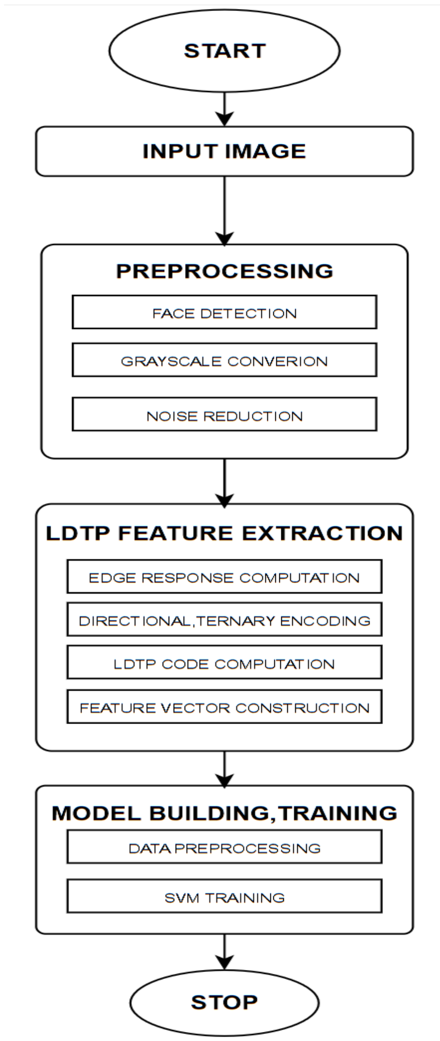

Fig. 1. Flow chart of Facial Expression Recognition.

A. Preprocessing and Face Detection

1) Face Detection: Facial expression recognition requires the accurate localization and normalization of the facial area to extract meaningful features. In the proposed system, each input image is first transformed into a grayscale image to decrease computational complexity and make the system colour independent. In image datasets where the facial area is not already pre-cropped, facial localization is carried out by a Haar Cascade classifier to locate the facial bounding box. The located ROI is then extracted and normalized to a fixed size.

2) Preprocessing: Preprocessing is a significant step in enhancing image quality, achieving uniform directionality, and ensuring stability in LDTP encoding. These factors are significant, particularly in the case of low-resolution facial images.

Initially the RGB face image is transformed to the grayscale image. This is done using the following formula as shown in (1)

$$I_{gray} = 0.2989 \cdot R + 0.5870 \cdot G + 0.1140 \cdot B \qquad (1)$$

where R, G, B are the red, green, and blue components.

Then a Contrast-Limited Adaptive Histogram Equalization (CLAHE) filter is used to enhance local contrast, bringing out small details of the facial muscles. It is essential to have clear-direction gradients in order to calculate the LDTP. This is done as shown in (2)

$$I_{enhanced} = \mathrm{CLAHE}(I_{gray}) \qquad (2)$$

Then, all the facial images will be resized into a fixed spatial resolution of 48x48 pixels to maintain uniformity within the data and match the dimensions required during the feature extraction pipeline. The resulting image is contrast-normalized using CLAHE to enhance facial muscle boundaries. Contrast normalization is essential to enhance the robustness of the overall directional response calculated during the Local Directional Ternary Pattern feature extraction process. The resulting normalized 48×48 grayscale facial image is now sent to the LDTP feature extraction block.

B. LDTP Feature Extraction

Facial expression recognition is based on the extraction of minute variations in facial expressions, which are mainly expressed as local directional texture changes. Conventional texture features like Local Binary Patterns (LBP) are not suitable for capturing directional gradient information of facial muscle activations, as they mainly focus on intensity variations [1], [2]. To overcome this drawback, the proposed method uses Local Directional Ternary Patterns (LDTP) as the feature descriptor because of their capability to capture directional sensitivity and noise robustness [4], [5].

1) Directional Gradient Computation**:** Facial expressions are distinguished by subtle and organized muscle movements, which appear as directional texture changes on the facial surface. These texture changes are not captured by analyzing pixel differences alone but by examining edge information in various directions. In the Local Directional Ternary Pattern (LDTP) scheme, edge information is derived from the computation of directional gradient responses that highlight expression-related facial structures.

For a normalized grayscale facial image I(x, y), directional gradients are determined by convolving the image with a set of compass masks, each of which is tailored to detect edges in a particular direction [4]. These masks help in the extraction of directional edge information related to facial muscle activities like eyebrow movement, lip extension, and cheek deformation. A conceptual representation of the compass masks employed for directional response computation is depicted in Fig. 2.

$$\underset{M_0}{\begin{bmatrix}-1&0&1\\-2&0&2\\-1&0&1\end{bmatrix}}\quad\underset{M_1}{\begin{bmatrix}0&1&2\\-1&0&1\\-2&-1&0\end{bmatrix}}\quad\underset{M_2}{\begin{bmatrix}1&2&1\\0&0&0\\-1&-2&-1\end{bmatrix}}\quad\underset{M_3}{\begin{bmatrix}2&1&0\\1&0&-1\\0&-1&-2\end{bmatrix}}$$

$$\underset{M_4}{\begin{bmatrix}1&0&-1\\2&0&-2\\1&0&-1\end{bmatrix}}\quad\underset{M_5}{\begin{bmatrix}0&-1&-2\\1&0&-1\\2&1&0\end{bmatrix}}\quad\underset{M_6}{\begin{bmatrix}-1&-2&-1\\0&0&0\\1&2&1\end{bmatrix}}\quad\underset{M_7}{\begin{bmatrix}-2&-1&0\\-1&0&1\\0&1&2\end{bmatrix}}$$

Fig. 2. Robinson Compass Masks Used for computing the directions

For pixel location (x, y), the directional response Ri (x, y) for the k-th compass mask is calculated as shown in (3)

$$R_k(x, y) = I(x, y) * M_k, \quad k=0,1,2,3,4,5,6,7 \tag{3}$$

Where M_k represents the k-th compass mask and * is the convolution operator. The magnitude of the directional responses is analyzed to determine the dominant orientations at each image point. By concentrating on the dominant directional gradients, LDTP helps in retaining significant facial texture information while discarding weak and noise-driven responses, thus laying a solid groundwork for ternary pattern construction.

2)Ternary Pattern Encoding: To improve robustness against illumination changes and noise gradients, LDTP uses a ternary encoding scheme rather than the binary thresholding method [4], [5]. This method allows the elimination of weak and insignificant directional information while retaining strong directional information related to facial muscle movements. For each directional response R_k, the ternary encoding scheme is defined as shown in (4)

$$T_k(x,y) = \begin{cases} +1, & R_k(x,y) \geq -\sigma \\ -1, & R_k(x,y) \leq +\sigma \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

Where σ is a threshold value used to distinguish strong directional information from weak or noise-induced gradients. Pixels with a zero value represent smooth areas and are not considered for further encoding. The ternary encoding method is very effective in improving noise robustness while retaining sensitivity to expression-related texture variations.

3)LDTP Code Construction : After the computation of the directional gradient and ternary mapping, the Local Directional Ternary Pattern (LDTP) code is formed to encode the primary directional information and edge polarity simultaneously for each pixel location. This enables the description of local facial texture variations caused by the expression-driven muscle contraction activity.

For each pixel location (x, y), the analysis of the directional response magnitudes is carried out to identify the dominant directions. The indices of the dominant and secondary directional responses are denoted by D1(x, y) and D2(x, y), respectively, and their ternary values are represented by T1(x, y) and

T2(x, y), respectively. The LDTP code at pixel location (x, y) is expressed in (5)

$$\mathrm{LDTP}(x,y) = D_1(x,y) \ll 6 + T_1(x,y) \ll 4 + D_2(x,y) \ll 2 + T_2(x,y) \quad (5)$$

where << represents the left bit-shift operator. This representation encodes the primary and secondary directional indices along with the polarity information simultaneously in a compact form. The pixels in smooth regions, represented by zero-valued ternary responses, are removed from the encoding process to avoid noise-driven gradients from affecting the descriptor.

Through the simultaneous encoding of the directional orientation and polarity information, the LDTP code map highlights the expression-related facial texture patterns and ignores the less important variations, thus establishing a robust local descriptor for the subsequent spatial aggregation of features.

4)Spatial Enhanced LDTP Representation*:* Facial expressions are spatial, where different facial parts contribute unevenly to the display of expressions [2], [8]. For example, the display of happiness is dominated by the mouth part, while the display of surprise is dominated by the eyebrow and eye parts. Therefore, it is not sufficient to use a global representation of texture to capture such localized facial muscle activities for expression display.

To capture spatial information, the normalized facial image is divided into several non-overlapping regions, and LDTP features are extracted independently for each region. This spatial representation allows the descriptor to capture region-wise directional texture information that is essential for expression distinction.

For each spatial region r, a histogram of LDTP codes is computed to describe the distribution of directional ternary patterns in the region. The regional histogram is defined as shown in (6)

$$h_r(k) = \#\{(x,y) \in r \mid \mathrm{LDTP}(x,y) = k\} \quad (6)$$

Where #(·) is the counting operation. The histograms of all spatial regions are concatenated to obtain a comprehensive feature vector that encodes both local texture information and spatial facial structure. By using this approach of representing spaces, the discriminative ability of the LDTP descriptor is significantly enhanced, where the localized expression-related texture information is preserved.

5)Feature Normalization: To make the feature representation scale-independent and suitable for kernel-based classifiers, the concatenated LDTP feature vector is normalized using L1 normalization as shown in (7)

$$\widehat{F_i} = \frac{F_i}{\sum_j F_j + \varepsilon} \quad (7)$$

Where ε is a small constant added to avoid division by zero. This then forms a normalized high dimensional LDTP feature vector that represents the distribution of directional ternary patterns on the face and is used as the LDTP component of the final feature representation.

III. EXPERIMENTAL SETUP

This section describes the dataset used for experimentation, data preparation and splitting, feature normalization, classifier configuration, and evaluation criteria used to assess the performance of the proposed LDTP-based facial expression recognition system.

A. Dataset Description

The proposed system is evaluated using a publicly available version of the Extended Cohn-Kanade (CK+) dataset, which is a widely used benchmarking dataset for facial expression recognition in a controlled setting [7]. The dataset contains facial images of seven categories of emotions: anger, contempt, disgust, fear, happiness, sadness, and surprise. Each image captures the peak expression of an emotion and is face-normalized to eliminate variations in pose and scale.

All facial images are converted to grayscale and resized to a uniform resolution of 48 × 48 pixels to facilitate compatibility with the LDTP feature extraction process.

Fig. 3. Sample facial images from CK+ Dataset

This figure provides a qualitative insight into the inter-class similarities and the minute variations in facial muscles that make facial expression recognition a challenging task.

B. Dataset Preparation and Stratified Splitting

Preparing the LDTP feature extraction dataset in a manner that facilitates reliable training and accurate evaluation of the classification model is the next step. Each facial image is described by a high-dimensional LDTP feature vector, as well as its corresponding emotion label. Mathematically, the dataset can be written as in (8), (9).

$$X = \{x_i\}_{i=1}^{N} \tag{8}$$

$$Y = \{y_i\}_{i=1}^{N} \tag{9}$$

Where $x_i \in R^d$ denotes the LDTP feature vector of the i^{th} sample and y_i , represents its associated emotion class.

To evaluate the generalization ability of the proposed facial expression recognition system, the dataset is split into training and testing sets. A stratified train-test split is used, where 80% of the instances from each emotion class are used for training, and the remaining 20% are held for testing. This split can be written as shown in (10)

$$X_{train}, X_{test}, Y_{train}, Y_{test} = \mathrm{Split}(X, Y, \rho) \tag{10}$$

Where ρ =0.2 denotes the proportion of test samples.

The need for stratification is especially prominent in facial expression recognition tasks, where the number of instances in each emotion class is likely to be imbalanced. For instance, happy or surprise expressions are likely to have many more instances than contempt or fear expressions. The stratified split ensures that the number of instances in each class is roughly preserved in both the training and testing sets as shown in (11).

$$P(y = k \mid Y_{train}) \approx P(y = k \mid Y_{test}), \forall k \tag{11}$$

This ensures that the classifier is trained and tested under similar conditions, thereby avoiding any potential biases that may arise from an imbalanced representation of certain emotions in the test set.

C. Feature Normalization and Label Encoding

As explained in Section II-B.5, LDTP histograms are normalized using L1 normalization as a preprocessing step during feature extraction to make the histogram representation scale-independent. Later, before training the classifiers, z-score normalization is carried out to normalize the feature vectors. The LDTP feature vectors derived using spatial histogram concatenation are naturally high-dimensional and can have large differences in the scale of the feature values across different feature dimensions. These differences can adversely affect the training of distance-based and margin-based classifiers, such as Support Vector Machines (SVMs). To counter this problem, feature normalization is carried out before classification. In this research, z-score normalization is used to normalize the features. To normalize a feature vector $x_i = [x_{i1}, x_{i2},, x_{id}]$ the following steps are carried out as shown in (12)

$$x'_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \tag{12}$$

Where, μ_j is the mean of the j^{th} feature computed from the training set, σ_j is the corresponding standard deviation.
This transformation in (13) ensure that each feature dimension has zero mean and unit variance.

$$E[x'_j] = 0, \quad VAR(x'_j) = 1 \tag{13}$$

Feature normalization is a crucial step in SVM-based classification because SVMs are sensitive to the relative scales of the input features [1]. Without feature normalization, features with larger scales in their numeric values could potentially overwhelm the optimization process, causing poor decision boundaries. By applying feature normalization, all LDTP feature values contribute equally to the classification decision, and this improves the training process.

To avoid leakage of information, the normalization parameters μ_j and σ_j are calculated solely on the basis of the training data and are then used as is on the test data.

Moreover, the categorical labels of the emotions are converted to numerical values using label encoding. Label encoding allows supervised learning to be performed efficiently. Label encoding ensures that the output of the classifier is interpreted in the same manner during both training and testing.

D. Classifier Configuration and Training

Finally, after the feature extraction and normalization process, the last step of the proposed facial expression recognition system is learning a decision model to relate the LDTP feature space to the corresponding classes of emotions. Due to the non-linear nature of the high-dimensional LDTP feature space, a Support Vector Machine (SVM) classifier with a radial basis function (RBF) kernel is used for classification [9], [10]. Let the normalized LDTP feature vectors and their corresponding labels in the training data be represented as in (14).

$$\{(x_i, y_i)\}_{i=1}^{N} \tag{14}$$

Where $x_i \in R^d$ represented the standardized LDTP feature vector of the i^{th} facial image and $y_i \in \{1, 2…., K\}$ represents its associated emotion class.
The SVM learns a non-linear decision function as shown in (15)

$$f(x) = \sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b \tag{15}$$

Where α_i are the support vector coefficients, b is the bias term, K denoted the RBF kernel defined as (16)

$$K(x_i, x_j) = \exp\left(-\gamma||x_i - x_j||^2\right. \tag{16}$$

In this case, the parameter γ is used to control the size of the Gaussian kernel and the effect of individual training samples. The regularization parameter C is used to control the trade-off between maximizing the margin and minimizing misclassification error. These parameters are chosen through empirical tuning and remain constant throughout the experiments for fair comparison.

In the facial expression recognition problem, the classes are imbalanced, meaning that some facial expressions are more common than others. To counter this problem, class-weight balancing is used during the training of the SVM, which gives higher misclassification costs to the minority classes, thus ensuring that all classes of facial expressions are balanced during training.

All parameters of the classifiers are determined solely based on the training data. Once the parameters have been determined, the trained model is tested on the held-out test data to determine the classification accuracy based on the chosen evaluation metric.
E. Performance Analysis

For the evaluation of the efficiency of the proposed system based on LDTP for facial emotion recognition, a number of performance metrics have been used. These performance metrics include Accuracy, Precision, Recall, and F1 Score. These performance metrics show a clear understanding of the efficiency of a classifier for predicting different classes.

1) Accuracy: It denotes the overall correctness of the model, which is represented as shown in (17)

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

Where,
TP (True Positives): correctly classified emotion samples,
TN (True Negatives): correctly rejected as other emotions,
FP (False Positives): incorrectly predicted as another emotion,
FN (False Negatives): incorrectly classified category.

In the case of facial expression recognition, accuracy is a high-level measure of system performance on all categories of emotion. However, because accuracy does not account for the different types of classification errors, it may not be a complete measure of performance on individual categories of emotion, especially in the case of class imbalance.

2) Precision: It calculates the model's dependability in its predictions about a particular emotion as shown in (18)

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{18}$$

A high precision measure of an emotion class means that most of the images classified into that class are actually correct. In facial expression recognition, precision is particularly important in applications where incorrect emotion detection could result in the misinterpretation of a subject's emotional state.

For instance, higher precision measures are typically obtained for expressions such as happy and surprise, which involve significant facial muscle activity and result in strong directional texture patterns. These distinctive features make the LDTP descriptor more capable of distinguishing these expressions.

3) Recall: This shows the ability of the model to recall all the examples associated with a particular emotion. It is calculated as shown in (19)

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{19}$$

High values of recall measures show that most of the images of a particular emotion have been successfully detected by the system. Low recall measures for some expressions, such as sad and angry, show that these expressions are more likely to be confused with other visually similar expressions. This can be attributed to the slight differences in the muscle activities of the face and the overlapping texture patterns, which are harder to extract using texture descriptors alone.

4) F1-Score: To achieve a balance between precision and recall, F1-score is used. It is calculated as shown in (20)

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{20}$$

This is particularly important in multi-class facial expression recognition tasks where there is class imbalance. For instance, classes such as disgust or contempt may contain fewer examples than happy classes. The F1 measure considers both false positives and false negatives, giving a fair evaluation of the performance of the classifiers on all classes of emotions.

5) Inter-Class Confusion Analysis: Though the proposed LDTP-based approach is quite effective in overall recognition accuracy, some inter-class confusions are observed, particularly among the emotions that have similar activations of facial muscles. Emotions such as happy and surprise, which have prominent activations of facial muscles in the mouth and eyebrow regions, are recognized with high accuracy due to the presence of strong and localized directional texture patterns. But the recall values for sad and angry expressions are lower since they are often confused with similar expressions like contempt and fear, respectively.

The reason for the confusions is the overlapping responses of directional edges in regions around the eyebrows and mouth, where the activations of facial muscles are small, resulting in similar texture patterns. As the proposed approach is based on texture descriptors, it is hard to distinguish between the expressions. But the spatial LDTP encoding method reduces these confusions by preserving the localized facial information, which enhances the discriminative capability of the texture descriptors.

IV. RESULTS AND DISCUSSION

This section presents the quantitative analysis of the proposed LDTP-based facial expression recognition system on the CK+ dataset. The recognition performance of the proposed system is presented in terms of overall recognition accuracy and class-wise analysis, followed by the discussion on the performance analysis and confusions among classes.

A. Overall Recognition Performance

The proposed LDTP-based facial expression recognition system performs well on the CK+ dataset

© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715104

with an image-level stratified train-test split. With an 80-20 split in training and testing, the system achieves an overall classification accuracy of 89.34%.

The results achieved clearly indicate that the combination of directional ternary encoding, spatially enhanced LDTP representation, and non-linear SVM classification is effective in extracting the discriminative facial texture patterns for different expressions on the CK+ dataset [4], [5], [6]. The use of spatially localized LDTP features enables the system to capture the subtle regional variations of facial muscle activity, thereby contributing to the improved recognition performance on different classes of expressions.

B. Class-wise Performance Evaluation

In order to comprehend the performance of the proposed framework, the performance analysis is conducted using precision, recall, and F1-score for each of the emotion classes. The performance analysis is helpful in comprehending the classification performance, especially in the case of class imbalance and similarity between classes.

Table I illustrates the class-wise precision, recall, and F1-score of the proposed LDTP-SVM framework for the seven facial emotion classes. Table II presents the overall performance analysis, including the classification accuracy and macro-averaged precision, recall, and F1-score.

Emotion	Precision	Recall	F1-Score	Support
Anger	0.92	0.89	0.91	27
Contempt	1.00	0.73	0.84	11
Disgust	0.97	0.91	0.94	35
Fear	1.00	0.67	0.80	15
Happy	0.91	0.98	0.94	42
Sadness	0.92	0.65	0.76	17
Surprise	0.80	1.00	0.89	50

Table II. Overall Performance Metrics

Overall Metric	Score
Accuracy	89.34%
Macro Average F1-Score	0.87
Weighted Average F1-Score	0.89

As seen in Table I, facial expressions involving strong localized activations of facial muscles, such as happiness and surprise, result in better recognition performance. This is because these facial expressions have strong directional edges around the mouth and eyebrow areas, which are well represented by the LDTP descriptor using directional ternary encoding and spatial feature aggregation. On the other hand, facial expressions like sadness and fear have relatively lower recall values, as these expressions involve weak facial deformations and are visually similar to other categories of facial expressions.

The performance metrics of the proposed LDTP-based facial expression recognition system are presented in Table II. Although overall accuracy is a measure of the correctness of classifications, macro-averaged F1-score is a more balanced measure of performance, giving equal weight to all classes of emotions irrespective of their occurrence. The weighted F1-score also takes into consideration the imbalance in classes by assigning weights to classes based on their occurrence in the dataset. The fact that the overall accuracy and weighted F1-score are in close agreement shows that the system is performing well on the classes of emotions that occur frequently, while the slight drop in macro-averaged F1-score shows that it is more difficult to classify the less frequent ones.

C. Comparative Evaluation with Baseline Models

For a complete assessment of the efficacy of the proposed LDTP-based framework, a comparative study is performed with popular handcrafted texture feature descriptors, including Local Binary Patterns (LBP), Local Directional Patterns (LDP), basic LDTP, and the proposed LDTP with spatial region separation [1], [4], [8]. These approaches naturally evolve from intensity-driven texture feature encoding to directionally and spatially advanced encoding.

Each of these approaches is tested in the same experimental setting, with the same preprocessing steps, feature normalization technique, classifier setup (RBF SVM), and image-level stratified train-test split. This helps in making a fair comparison, where the differences in performance can be ascribed to the feature descriptor alone.

© MAR 2026 | IRE Journals | Volume 9 Issue 9 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I9-1715104

Table III. Comparative Evaluation Od Texture Based Fer Models

Model	Feature Description	Accuracy (%)
M1	LBP (Local Binary Pattern)	51.7
M2	LDP (Local Directional Pattern)	53.89
M3	Baseline LDTP	78.10
M4 (Proposed)	LDTP + Spatial Regions	89.34

As seen from Table III, it is clear that the addition of directional information has a significant positive impact on recognition performance, as seen when comparing the results of LBP to LDP-based features. The addition of ternary encoding in the baseline LDTP further improves robustness by eliminating the effect of weak and noise-driven gradient responses, thereby improving classification accuracy and macro F1-score.

The proposed LDTP framework incorporating spatial region separation performs the best on all metrics of evaluation [5], [6]. This is significant because it emphasizes the need to retain localized facial texture information, as different facial regions do not contribute equally to the expression formation process. The proposed method is able to effectively capture region-specific texture variations not captured by global features.

The comparative analysis clearly shows that the proposed spatially enhanced LDTP framework offers a more discriminative and robust representation for facial expression recognition than traditional handcrafted texture features.

D. Class-wise Confusion Analysis

Fig. 4 shows the confusion matrix for the proposed Robinson-based spatial LDTP model on the CK+ database. The confusion matrix is dominated by its diagonal, which reflects the correct recognition of most categories of emotions. The small amount of confusion is mainly between expressions that are close in appearance, such as angry and fear, and happy and sad, which involve overlapping activations of facial muscles around the eyes and mouth. These are expected difficulties in facial expression recognition and are not significant in number, reflecting the success of the proposed spatial LDTP model in capturing the discriminative texture information of the face.

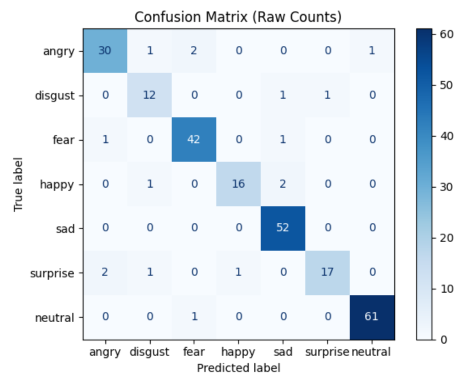

Fig. 4. Confusion matrix of proposed LDTP model

V. CONCLUSION

This paper has proposed an improved facial expression recognition system using the Local Directional Ternary Pattern (LDTP) technique, which aims to efficiently capture the discriminative facial texture patterns while being computationally efficient. By exploiting the directional ternary pattern and adding the spatial region separation, the proposed system improves the localized facial muscle movement representation, which is essential for distinguishing among various facial expressions.

The experimental results on the CK+ facial expression database with an image-level stratified train-test split show that the proposed LDTP-based system outperforms the conventional handcrafted texture feature descriptors, such as LBP, LDP, and basic LDTP representation [1], [4], [5], [6], [7]. The experimental comparison shows that the ternary pattern enhances the noise and illumination robustness, and the spatially enhanced LDTP representation further improves the discriminative representation by preserving the region-specific texture information on the facial area.

The experimental results show that facial expressions with high localized muscle activity, such as happiness and surprise, are recognized more accurately, while facial expressions involving subtle facial deformations are more difficult to recognize due to overlapping texture patterns. However, the proposed spatial LDTP-based system successfully reduces these confusions by preserving the localized directional information, resulting in a more balanced recognition performance among various facial expressions.

In conclusion, the proposed method provides a robust and efficient hand-crafted solution for facial expression recognition, which is highly suitable for scenarios with limited training data and computational capabilities. Future research work may include exploring subject-independent evaluation protocols, adaptive spatial partitioning schemes, and hybrid models that integrate the proposed method with deep learning models.

REFERENCES

[1] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[2] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," Image and Vision Computing, vol. 27, no. 6, pp. 803–816, 2009.

[3] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, Computer Vision Using Local Binary Patterns, Springer, London, 2011.

[4] K. E. Chahi, Y. Ruichek, J. Mejdoub, and R. Touhami, "Local directional ternary pattern: A new texture descriptor for texture classification," Computer Vision and Image Understanding, vol. 169, pp. 14–27, Apr. 2018.

[5] S. R. Ryu, J.-H. Park, and Y. M. Ro, "Local directional ternary pattern for facial expression recognition," IEEE Transactions on Image Processing, vol. 26, no. 12, pp. 6006–6021, Dec. 2017.

[6] Y. Zhang, Z. Jiang, J. Li, J. Sun, and Z. He, "Multi-feature fusion local directional ternary pattern for facial expression recognition," Alexandria Engineering Journal, vol. 61, no. 12, pp. 11945–11958, Dec. 2022.

[7] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The Extended Cohn–Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2010, pp. 94–101.

[8] G. Zhao and M. Pietikäinen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 6, pp. 915–928, 2007.

[9] S. L. Happy and A. Routray, "Automatic facial expression recognition using features of salient facial patches," IEEE Transactions on Affective Computing, vol. 6, no. 1, pp. 1–12, 2015.

[10] P. Dhivya and S. V. Kumar, "Facial expression recognition for low resolution images using LBP, HOG, SIFT and optical flow with multiclass SVM," SSRG International Journal of Electrical and Electronics Engineering, vol. 7, no. 3, 2020.

[11] S. Li and W. Deng, "Deep facial expression recognition: A survey," IEEE Transactions on Affective Computing, vol. 13, no. 3, pp. 1195–1215, Jul.–Sep. 2022.

[12] A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," in Proc. IEEE Winter Conference on Applications of Computer Vision (WACV), 2016.

[13] P. Burkert, F. Trier, M. Z. Afzal, A. Dengel, and M. Liwicki, "DeXpression: Deep convolutional neural network for expression recognition," Pattern Recognition Letters, vol. 34, no. 6, pp. 825–830, 2013.

[14] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 1, pp. 39–58, 2009.

[15] A. A. A. Abdelmaksoud, M. Nabil, and I. Farag, "Advances in facial expression recognition: A survey of methods, benchmarks, models, and datasets," Information, vol. 15, no. 3, Art. no. 135, Feb. 2024.