IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, May 2026
Impact Factor: 8.2

An Effective Facial Emotion Recognition Framework Using Advanced Feature Extraction and Hybrid Feature Fusion Techniques

T. J Saravanan
Assistant Professor, Department of Information Technology
K L.N. College of Engineering, Sivaganga, India
tjsar888@gmail.com

Abstract: *Emotions play a vital role in human communication and contribute significantly to building a healthy and socially connected society. In recent years, researchers have developed various Facial Expression Recognition (FER) systems to interpret human emotions as the demand for advanced human–computer interaction applications continues to grow. However, creating an FER system that performs effectively in real-world environments remains a challenging task due to variations in lighting, pose, and other uncontrolled imaging conditions.*

This research work presents an FER system that utilizes both machine learning and deep learning techniques to recognize emotions from static images as well as video sequences. Within the machine learning framework, two novel methods are introduced: the Statistical Shape Projection Model (SSPM) and the Enhanced Multi-Feature Fusion Model (EMFM), which focus on geometric and appearance-based feature extraction methods respectively. In the deep learning framework, a Maximum Boosted Convolutional Neural Network (MBCNN) and a hybrid architecture combining MBCNN with Long Short-Term Memory (LSTM) are proposed to improve the accuracy of emotion classification.

The Statistical Shape Projection Model (SSPM) achieves high classification performance by extracting meaningful features using a statistical shape model combined with integral projection analysis. In this approach, a static image is taken as input and its edge information is enhanced through the use of un sharp masking. The statistical shape model adapts the new input image to the trained model structure, allowing the deformation patterns to be learned efficiently. The effectiveness of the proposed feature extraction method is further improved when combined with projection-based analysis. Additionally, an edge enhancement technique highlights prominent structural details, enabling the system to identify discriminative features more effectively for accurate facial expression recognition.

Keywords: Facial Emotion Recognition, Computer Vision, Feature Extraction, Machine Learning, Human Computer Interaction.

I. INTRODUCTION

Emotions play an essential role in everyday human communication. Emotion can be described as a complex psychological state that involves behavioral, physiological, and cognitive responses. During interpersonal interaction, individuals can naturally observe changes in human behavior that reflect emotional states. Understanding and interpreting these emotional variations can help improve interpersonal relationships and contribute to building a healthier and more harmonious society. Humans are capable of recognizing emotions through several modalities such as facial expressions, speech, tone of voice, body gestures, and posture, as well as through physiological signals captured by biosensors including Electromyography (EMG), Electroencephalography (EEG), and Electrocardiography

DOI: 10.48175/IJARSCT-35106

(ECG). Among these modalities, facial expressions are one of the most natural and immediate ways of conveying a person's feelings, intentions, and attitudes.

To improve automatic emotion recognition, researchers have developed Facial Expression Recognition (FER) systems that attempt to model aspects of human emotional intelligence. Due to its wide range of practical applications, FER has become an important research area within computer vision and Human–Computer Interaction (HCI).

Human–Computer Interaction focuses on designing computer systems that can understand and respond to human behavior. It incorporates computational methods related to learning, perception, and reasoning to simulate aspects of human intelligence (Breazeal et al., 2003). Whether interaction occurs directly or indirectly, facial emotions are often communicated effortlessly between individuals. Therefore, this research emphasizes the use of facial expressions as a primary indicator for emotion detection.

Facial expressions are produced by the movement of facial muscles beneath the skin. These movements cause variations in different facial regions, forming distinct patterns based on features such as the eyes, nose, mouth, eyebrows, and cheeks. By learning the characteristics of these patterns, a computer system can classify emotions based on observed facial expressions. Ultimately, the number and quality of informative features extracted within the feature space significantly influence the accuracy of the emotion classification process.

1.1.1 Overview of Facial Expression Recognition

Early studies on facial expression analysis were introduced by Darwin (1872), which later gained significant attention from researchers. According to Mehrabian (1974), approximately 55% of communication is conveyed through facial expressions, 38% through vocal tone and verbal cues, and only 7% through spoken words. Compared to other modalities, facial expressions contain rich and multidimensional information.

Psychological studies classify emotions into three main approaches: discrete emotion categories, dimensional models, and appraisal-based approaches. Among these, Ekman and Friesen (1971) proposed six universal facial expressions representing basic emotional categories: Anger, Disgust, Fear, Happiness, Sadness, and Surprise. In this research work, the categorical approach is primarily adopted by learning distinctive features from facial expressions. Each emotion corresponds to a specific facial pattern that can be recognized consistently across different cultures.

The six basic facial expressions and their typical facial changes are described as follows:

- Happiness: This expression is associated with smiling and joy. The eyebrows appear relaxed, the mouth widens, and wrinkles often form at the outer corners of the eyes.
- Sadness: This expression reflects sorrow or unhappiness. The inner portions of the eyebrows rise upward, and the corners of the lips move downward.
- Fear: Fear indicates anxiety or alarm. The eyebrows lift and move closer together while the eyes open wider than usual.
- Anger: In this expression, the inner eyebrows move downward and closer together, and the lips may tighten or expose the teeth.
- Surprise: Surprise can occur in both positive and negative situations. It is characterized by raised eyebrows, widened eyes, and an open mouth.
- Disgust: This expression typically occurs when encountering unpleasant tastes or odors. The nose wrinkles, and the upper lip lifts or curls upward.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, May 2026
Impact Factor: 8.2

Humans are naturally capable of identifying emotions, but recognizing emotions accurately in application-based environments is often challenging. To address this issue, intelligent systems have been developed to enable machines to recognize emotions efficiently and sometimes even more consistently than humans. Over the past few decades, automated **Facial Expression Recognition (FER)** systems have been developed using various **machine learning techniques** to classify different emotional states. Many researchers have focused on emotion recognition through facial expressions because facial cues provide immediate emotional information without requiring additional sensors or external input. This capability is highly valuable for improving machine learning models and enabling computers to interact with humans in a more natural and intuitive manner, similar to human-to-human communication.

Although machines are designed to perform tasks related to human behavior, several challenges still remain. These include handling **large-scale datasets, improving processing speed, increasing recognition accuracy, and reducing computational complexity**. As a result, there is significant scope for developing more efficient machine learning algorithms to support a wide range of **Human–Computer Interaction (HCI)** applications.

In recent years, **deep learning techniques** have gained significant attention in the development of FER systems and have demonstrated superior performance in image classification tasks. With the availability of large datasets, powerful computational resources such as **Graphics Processing Units (GPUs)**, and increased memory capacity, deep learning methods have addressed many challenges in fields such as **computer vision, human–computer interaction, and biomedical analysis**. This research work aims to design an FER system capable of automatically recognizing human emotions from **static images and video sequences** based on facial expressions. Such emotion recognition systems have numerous practical applications, including **psychological studies, patient monitoring systems, video surveillance, face recognition, and cognitive behavior analysis**.

II. LITERATURE REVIEW

Several studies have been conducted in the field of **facial expression recognition (FER)** to improve the accuracy of emotion detection using computer vision and machine learning techniques. Early research in this area focused mainly on **traditional machine learning methods** that relied on handcrafted feature extraction techniques such as **Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), and Gabor filters**. These methods extracted important facial features and used classifiers like **Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Artificial Neural Networks (ANN)** to categorize human emotions.

Later, researchers introduced **geometric-based and appearance-based approaches** to analyze facial expressions. Geometric methods focused on the movement and position of facial landmarks such as the eyes, eyebrows, nose, and mouth, while appearance-based methods extracted texture features from facial images. These approaches improved the performance of FER systems but were still affected by variations in lighting conditions, head pose, and facial occlusions.

With the advancement of computational power and availability of large datasets, **deep learning techniques** have become widely used for facial expression recognition. **Convolutional Neural Networks (CNNs)** are commonly applied for automatic feature extraction and emotion classification from facial images. CNN-based models have shown

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 4, May 2026

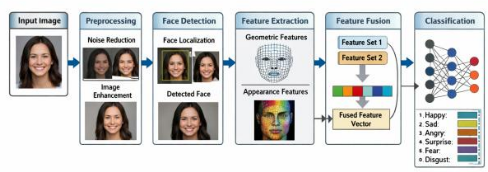
Impact Factor: 8.2

significant improvements in accuracy compared to traditional machine learning methods. In addition, hybrid models combining **CNN with Recurrent Neural Networks (RNN) or Long Short-Term Memory (LSTM)** networks have been developed to analyze both spatial and temporal information from video sequences.

Recent studies focus on improving FER systems by integrating **multiple feature extraction techniques, deep learning architectures, and multimodal data sources**. These advancements help to enhance emotion recognition performance and enable the development of intelligent systems capable of interacting with humans in a more natural and effective manner.

III. SYSTEM ARCHITECTURE

The proposed facial emotion recognition system consists of multiple stages including preprocessing, face detection, feature extraction, feature fusion, and classification. First, input facial images are preprocessed to improve image quality and remove noise. Next, face detection algorithms identify the facial region within the image. After detecting the face, important facial features are extracted using the proposed methods. Feature fusion techniques combine multiple descriptors to generate a discriminative feature vector. Finally, a machine learning classifier predicts the emotional category.



IV. PREPROCESSING AND FACE DETECTION

Preprocessing is an essential stage in a facial emotion recognition system, as it improves the quality of input images and prepares them for accurate feature extraction. Raw facial images often contain noise, illumination variations, and other distortions that can negatively affect recognition performance. Therefore, several preprocessing techniques are applied to standardize and enhance the images before further analysis.

Image Normalization

Image normalization is performed to maintain consistency across the dataset. Images are resized to a fixed dimension (for example, 128×128 or 256×256 pixels) and pixel intensity values are normalized to a specific range. This process ensures that all images have a uniform scale and reduces computational complexity during feature extraction.

Brightness Correction

Variations in lighting conditions can significantly affect the appearance of facial features. Brightness correction techniques adjust the illumination levels of the image so that facial features become more distinguishable. This step reduces the impact of uneven lighting and shadows.

Contrast Enhancement

Contrast enhancement methods such as **Histogram Equalization (HE) or Adaptive Histogram Equalization (AHE)** are used to improve the visibility of facial structures. These techniques increase the contrast between different regions of the image, allowing important facial details such as edges and wrinkles to become more prominent.

Grayscale Conversion

Color images contain three channels (RGB), which increase computational cost. Since most facial expression information lies in texture and intensity variations, color images are converted into **grayscale images**. This reduces the data dimensionality while preserving essential facial features required for emotion recognition.

DOI: 10.48175/IJARSCT-35106

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, May 2026
Impact Factor: 8.2

Face Detection

After preprocessing, the next step is to detect the location of the face within the input image. Face detection isolates the facial region from the background and eliminates irrelevant information.

Several well-known algorithms are used for face detection:

Haar Cascade Classifier

The **Haar Cascade classifier**, introduced by Viola and Jones, is one of the most widely used real-time face detection methods. It works by extracting Haar-like features and using an **AdaBoost classifier** to select the most relevant features. A cascade structure of classifiers quickly rejects non-face regions, making the method computationally efficient.

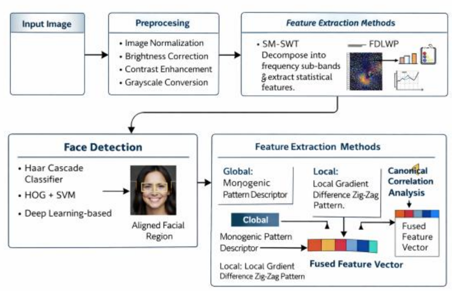
Histogram of Oriented Gradients (HOG) with Support Vector Machine (SVM)

The **HOG descriptor** captures edge and gradient orientation information from the image. These features describe the shape and structure of objects in the image. When combined with a **Support Vector Machine (SVM)** classifier, the HOG-based detector can accurately distinguish facial regions from the background.

Deep Learning-Based Face Detectors

Recent approaches employ deep learning models such as **Convolutional Neural Networks (CNNs)** for face detection. Models such as **MTCNN (Multi-task Cascaded Convolutional Networks)** and **Faster R-CNN** provide higher accuracy and robustness under challenging conditions such as varying illumination, occlusion, and head pose.

Once the face is detected, the facial region is cropped from the image. Facial alignment is then performed using **eye coordinates** as reference points. Aligning faces ensures that facial components such as the eyes, nose, and mouth are positioned consistently across all images, which improves the reliability of feature extraction.



V. FEATURE EXTRACTION METHODS

Feature extraction is a crucial step in facial emotion recognition, as it identifies meaningful characteristics from facial images that represent emotional expressions. The goal of feature extraction is to convert raw image data into a set of discriminative features that can be used for classification.

In this research, several feature extraction techniques are utilized to capture both **global and local facial characteristics.**

VI. EMOTION CLASSIFICATION

Classification algorithms assign emotion labels to extracted feature vectors. Several machine learning techniques can be used including Support Vector Machines, Artificial Neural Networks, Random Forest, and K-Nearest Neighbor. The classifier is trained using labeled datasets containing various facial expressions. During testing, the trained model predicts the emotional category for unseen facial images.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, May 2026
Impact Factor: 8.2

VII. EXPERIMENTAL RESULTS

The proposed system was evaluated using several benchmark datasets including JAFFE, CK+, MMI, RAF, Oulu-Cassia, SFEW, and LIRIS-CSE.

Performance was measured using accuracy, precision, recall, sensitivity, specificity, and F1-score.

Experimental results show that the proposed feature extraction and fusion techniques significantly improve recognition accuracy compared to traditional methods.

VIII. CONCLUSION

This paper presented an effective facial emotion recognition framework using advanced feature extraction and hybrid feature fusion techniques.

The experimental results demonstrate that the proposed approach achieves improved accuracy and robustness across multiple datasets.

Future work will focus on integrating deep learning architectures and developing real-time emotion recognition systems for practical applications such as healthcare monitoring and smart learning environments.

REFERENCES

[1]. Erfan Zangeneh & Aref Moradi 2018, 'Facial expression recognition by using differential geometric features', The Imaging Science Journal.

[2]. Fan, J, Xu, W, Wu, Y & Gong, Y 2010, 'Human tracking using convolutional neural networks', IEEE Transactions on Neural Networks, vol. 21, no. 10, pp. 1610-1623.

[3]. Fathallah, A, Abdi, L, Douik, A 2017, 'Facial expression recognition via deep learning', Proceedings of International Conference on Computer Systems and Applications, Hammamet, pp. 745-750.

[4]. Fatima, ZS, Abdellah, M & Mohamed, K 2018, 'Emotion Recognition from Facial Expression based on Fiducial points detection and using neural network', International Journal of Electrical and Computer Engineering, vol. 8, no. 1, pp. 52-59.

[5]. Garcia-Mateos, G, Ruiz, A, Lopez-de-Teruel, P E 2002, 'Face Detection using Integral Projection Models', Lecture Notes in Computer Science, Structural, syntactic, and statistical Pattern Recognition, vol. 2396. https://doi.org/10.1007/3-540-70659-3_67.

[6]. Ghimire, D & Lee, J 2013, 'Geometric feature-based facial expression recognition in image sequences using multi-class adaboost and support vector machines', Sensors, vol. 13, no. 6, pp. 7714-7734.

[7]. 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, vol. 76, pp. 7803-7821.

[8]. Graves, A, Mayer, C, Wimmer, M & Schmidhuber, J 2008, 'Facial expression recognition with recurrent neural networks', Proceedings of Cognition for Technical Systems Conference, Santorini, Greece, pp. 1-6.

[9]. Haitham F Al-Mubarak 2014, 'Human face identification from profile projection', Al-Nahrain Journal of Science, vol. 1, no. 3, pp. 204-209.