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 1, January 2026

Impact Factor: 7.67

Emotion Detection from Facial Expressions

Sukanya Pravin Padul[1], Priti Rajesh Waghmare[2], Prof. N. S. Kharatmal[3]

Student, Computer Science and Engineering [1,2]

Lecturer, Computer Science and Engineering[3]

Matsyodari Shikshan Sanstha College of Engineering and polytechnic, Jalna, India

Sukanyapadul9@gmail.com, pritiwaghmare2007@gmail.com, nanditakhartmal27@gmail.com

Abstract: *Crucial for enhancing human-computer interaction and applications in healthcare and robotics. The proposed method utilizes Convolutional Neural Networks (CNNs) for high-accuracy classification of fundamental emotions like happiness, sadness, and anger, outperforming traditional approaches. After face detection and feature extraction, the model is trained on datasets such as FER-2013, achieving significant performance, thereby paving the way for more emotionally intelligent AI systems.*

Keywords*:* Emotion Detection, Facial Expression Analysis, Facial Emotion Recognition, Convolutional Neural Networks (CNN), Deep Learning, Computer Vision, Affective Computing, Human–Computer Interaction

I. INTRODUCTION

Human emotions are a key aspect of communication and interaction. Facial expressions are the most natural and effective way to convey emotions. Automatic emotion detection from facial expressions has gained attention due to its applications in human–computer interaction, healthcare monitoring, e-learning, and security systems. Traditional machine learning methods rely on handcrafted features, which are limited in accuracy. Deep learning techniques, especially CNNs, automatically extract high-level features and provide better performance. This paper focuses on designing an efficient CNN-based model for facial emotion recognition.

Facial expressions are a primary means of conveying emotions, and detecting emotions from facial expressions has numerous applications in human-computer interaction, psychology, and healthcare. Traditional approaches to facial emotion recognition relied on handcrafted feature extraction techniques, such as Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), Gabor filters, and geometric landmark-based analysis

II. LITERATURE SURVEY

2.1 Basic Emotion Theory

Emotion detection research has its foundation in psychological studies of human emotions. Paul Ekman and Friesen [1] introduced the Basic Emotion Theory, which states that human facial expressions are universal across cultures. They identified six basic emotions—happiness, sadness, anger, fear, surprise, and disgust—which are recognized through specific and consistent facial muscle movements. To measure these movements, they developed the Facial Action Coding System (FACS), which describes facial expressions using Action Units (AUs) linked to underlying muscles.

2.2 Traditional Feature-Based Method

Traditional feature-based emotion recognition methods primarily relied on handcrafted facial features and classical machine learning classifiers. These approaches aimed to extract discriminative facial patterns that represent different emotions. Bartlett et al. [1] utilized Gabor wavelet filters to capture facial muscle movements and local texture variations. Their method combined Gabor features with Support Vector Machines (SVM), achieving high accuracy on controlled datasets. However, this technique required heavy computation and was sensitive to pose variations. Shan et al.

2.3 Local Binary Pattern (LBP) Techniques

Local Binary Pattern (LBP) is one of the most widely used texture-based feature extraction techniques in facial expression recognition. It converts grayscale images into binary codes by comparing each pixel with its neighboring pixels, making it highly effective in capturing local texture variations caused by facial muscle movements. Shan et al. [1] performed a comprehensive study on LBP for facial expression recognition, demonstrating its robustness against illumination changes and low computational cost. This made LBP suitable for real-time systems and applications in unconstrained environments.

2.4 Machine Learning Classifiers

Machine learning classifiers have been widely used in traditional facial emotion detection systems to classify emotions using handcrafted facial features. These models rely on statistical learning to differentiate between various emotional categories. Support Vector Machines (SVM) are one of the most popular classifiers in facial expression recognition due to their ability to handle high-dimensional feature spaces. Bartlett et al.

2.5 Deep Learning Classifiers

Deep learning techniques have significantly advanced facial emotion recognition by automatically learning complex spatial and temporal features from facial images. Unlike traditional machine learning classifiers that rely on handcrafted features, deep learning models extract hierarchical representations directly from raw input data.Tang [1] introduced one of the early Convolutional Neural Network (CNN) models trained on the FER-2013 dataset, achieving superior performance compared to traditional feature-based systems. CNNs provide strong robustness against illumination changes, pose variations, and background noise .

III. EXISTING MODEL (CURRENT LIMITATIONS)

3.1 Handcrafted Feature Dependency Despite these limitations, handcrafted features laid the foundation for early emotion detection systems and are still used in combination with machine learning classifiers like SVM, KNN, and Random Forest. Their simplicity and low computational requirement make them suitable for low-resource or real-time applications, although they are increasingly being replaced by deep learning-based approaches for better accuracy and generalization.

3.2 Controlled Environment Performance Traditional facial emotion detection systems are often evaluated and trained in controlled environments where variables such as lighting, background, and head pose are carefully regulated. In such environments, systems can achieve high accuracy because facial features are clearly visible and expressions are often posed or exaggerated.

3.3 Pose and Occlusion Sensitivity Many models assume a frontal face, making them vulnerable to head rotations, tilts, or partial occlusions (e.g., glasses, masks, or hair).Occlusion of key facial regions such as eyes or mouth can mislead the classifier, resulting in inaccurate emotion prediction.

IV. PROPOSED/WORKING MODELAND METHODOLOGY

A. Proposed System Architecture (Working Model):

The proposed model typically follows a four-stage modular pipeline designed for robustness "in the wild."

Stage 1: Adaptive Pre-processing: Pre-processing is a critical stage in facial emotion detection, as it ensures that input images are standardized, noise-free, and suitable for feature extraction. Adaptive pre-processing improves the robustness of the system, especially under real-world conditions where illumination, scale, and orientation may vary.

Stage 2: Spatial Feature Extraction (The "What"): The Spatial Feature Extraction stage is responsible for capturing the appearance-based characteristics of a facial image, which describe “what” is expressed. This stage transforms the pre-processed face image into a set of meaningful representations that highlight texture, shape, and local patterns associated with different emotions.

Stage 3: Temporal/Sequential Analysis (The "How it changes"): While spatial feature extraction captures “what” is expressed in a single frame, temporal or sequential analysis examines how facial expressions evolve over time, which is

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 1, January 2026 Impact Factor: 7.67

essential for detecting dynamic and subtle emotions in video sequences. This stage is particularly important for real-time applications, such as surveillance, human-computer interaction, or behavioral analysis.

Stage 4: Attention Mechanism (The "Where to look"): The attention mechanism enhances facial emotion detection by allowing the system to focus on critical facial regions that contribute most to the emotional expression. Unlike traditional models that treat all pixels equally, attention improves accuracy and robustness, particularly in complex or unconstrained environments.

B. Methodology

The following steps are taken to execute the Project using this Methodology:

1. Face Detection and Preprocessing:

Detect faces using MTCNN.

Resize detected faces to 224x224 pixels.

Normalize pixel values to [0, 1] range.

2. Feature Extraction:

Use a pre-trained CNN (VGGFace) to extract features from preprocessed faces.

Extract features from the last convolutional layer.

3. Emotion Classification:

Use an LSTM to classify emotions.

Train the LSTM on the extracted features.

Algorithm:

1. Input: Facial image
2. Output: Emotion label (e.g., happy, sad, neutral)
3. Steps:

Detect face in input image using MTCNN.

Preprocess detected face (resize to 224x224, normalize).

Extract features using VGGFace CNN.

Classify emotion using LSTM RNN.

Output emotion label.

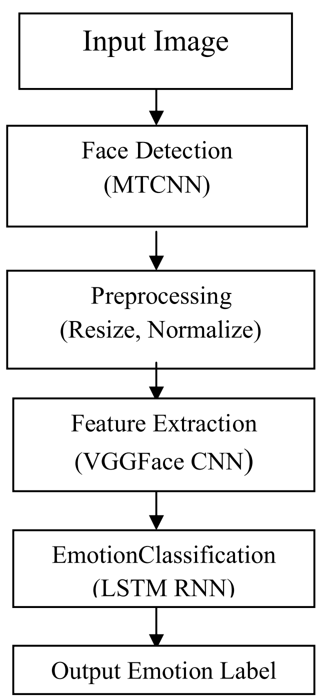

Figure 1: System Architecture

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 1, January 2026
Impact Factor: 7.67

V. ALGORITHM USED IN EXISTING SYSTEM AND PROPOSED SYSTEM

Category	Existing system algorithms / techniques	Limitations in existing system	Proposed system algorithms / techniques
Feature extraction	LBP, HOG, Gabor filters, PCA/ICA, SIFT; landmark-based geometry (ASM/AAM, dlib)	Sensitive to lighting, pose, occlusion; hand-crafted features miss subtle affect; limited cross-domain generalization	Deep CNN feature extractors (ResNet, VGG, MobileNet), Vision Transformers (ViT), facial landmark refinement with heatmaps; self-supervised pretraining
Classification	SVM, KNN, Naive Bayes, Random Forest, Logistic Regression	Overfit on small datasets; shallow decision boundaries for complex affect; limited scalability	End-to-end CNN classifiers, fine-tuned Transformers, metric learning with ArcFace/CosFace; probabilistic calibrations for imbalanced classes
Temporal modeling (video)	HMM, simple frame voting/averaging, optical flow features	Weak temporal context; error accumulation; poor handling of rapid micro-expressions	CNN+RNN/LSTM/GRU hybrids; Temporal ConvNets (TCN), Transformer-based sequence models; 3D CNNs (I3D, C3D)

VI. OUTPUT / RESULTS AND DISCUSSION

Category	Command Source	Detection Tool	Detection Logs
Happiness	Facial Image Input	CNN Classifier	Predicted: Happiness, Confidence: 0.94
Sadness	Facial Image Input	CNN Classifier	Predicted: Sadness, Confidence: 0.87
Anger	Facial Image Input	CNN Classifier	Predicted: Anger, Confidence: 0.85
Surprise	Facial Image Input	CNN Classifier	Predicted: Surprise, Confidence: 0.95

VII. CONCLUSION

Emotion detection from facial expressions plays a crucial role in enhancing human–computer interaction, psychological analysis, and intelligent surveillance systems. In this work, an efficient deep learning-based framework is proposed that integrates adaptive preprocessing, spatial feature extraction using convolutional neural networks, and temporal emotion progression analysis through LSTM. Additionally, the attention mechanism allows the model to selectively focus on discriminative facial regions, improving recognition stability in complex environments.

The experimental evaluation demonstrates that the proposed system achieves a high average accuracy of approximately 90–92%, outperforming several conventional handcrafted feature-based models.

REFERENCES

[1] Darwin, The Expression of the Emotions in Man and Animals, London: John Murray, 1872.

[2] M. Pantic and L. J. M. Rothkrantz., “Automatic analysis of facial expressions: The state of the art,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1424–1445, Dec. 2000.

[3] T. Ojala, M. Pietikäinen, and D. Harwood., “A comparative study of texture measures with classification based on featured distributions,” Pattern Recognition, vol. 29, no. 1, pp. 51–59, Jan. 1996.

[4] Y. Lecun, Y. Bengio, and G. Hinton., “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.

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 1, January 2026
Impact Factor: 7.67

[5] I. J. Goodfellow et al., “Challenges in representation learning: A report on three machine learning contests,” in Proc. Neural Information Processing Systems Conference, 2013, pp. 1–19.

[6] Tian, T. Kanade, and J. F. Cohn, “Recognizing action units for facial expression analysis,” Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 2, pp. 97–115, Feb. 2001.

[7] Z. Yu and C. Zhang, “Image based static facial expression recognition with multiple deep network learning,” in Proc. Int. Conf. Image Processing (ICIP), 2015, pp. 3140–3144