

# BUILDING HUMAN FACIAL EMOTIONS DETECTION MODEL

<sup>1</sup> J.R. Arun Kumar, <sup>2</sup> Vivek Jangid, <sup>3</sup> Nisha, <sup>4</sup> Virendra Yadav, <sup>5</sup> Monu

<sup>1</sup> Professor, Department of CSE, Modern Institute of Technology and Research Centre, Alwar, Rajasthan, India.

<sup>2,3,4,5</sup> UG Student, Department of AI & DS, Modern Institute of Technology and Research Centre, Alwar, Rajasthan, India

## Article Information

Received : 26 March 2026

Revised : 26 March 2026

Accepted : 27 March 2026

Published : 29 March 2026

*Corresponding Author:*

*Vivek Jangid*

**Abstract**— Facial emotion recognition plays an important role in enabling intelligent systems to understand human emotions and enhance human computer interaction. This paper presents a deep learning based facial emotion detection system capable of identifying multiple emotional states from facial images. The model is trained on a standard facial expression dataset containing classes such as happy, sad, angry, surprise, fear, disgust, and neutral. A convolutional neural network is used to automatically extract features from preprocessed images, where preprocessing includes resizing, normalization, and augmentation to improve performance. The model is evaluated on test data to ensure reliable emotion classification. The system is deployed as a user friendly web application that supports both image upload and real time webcam input. It provides predicted emotions with confidence scores and includes an audio feedback feature for better user interaction. The proposed system offers an effective solution for real time facial emotion recognition with applications in intelligent systems and human computer interaction. Future work focuses on improving accuracy and enhancing system capabilities for real world use.

**Keywords:** *Audio Feedback; Convolutional Neural Network; Deep Learning; Facial Emotion Recognition; Image Preprocessing; Real-Time Detection; Streamlit Application; Webcam Integration*

**Copyright © 2026: J.R. Arun Kumar, Vivek Jangid, Nisha, Virendra Yadav, Monu.** This is an open access distribution, and reproduction in any medium, provided Access article distributed under the Creative Commons Attribution License the original work is properly cited License, which permits unrestricted use.

**Citation: J.R. Arun Kumar, Vivek Jangid, Nisha, Virendra Yadav, Monu.** “BUILDING HUMAN FACIAL EMOTIONS DETECTION MODEL”, Journal of Science, Computing and Engineering Research, 9(3), March 2026.

## I. INTRODUCTION

Facial expressions are one of the most natural and powerful ways through which humans convey emotions. The ability to automatically recognize these emotions using computational methods has gained significant attention in the fields of computer vision, artificial intelligence, and human computer interaction. Facial emotion recognition systems aim to identify human emotional states such as happiness, sadness, anger, surprise, fear, disgust, and neutrality by analyzing facial features. These systems have wide applications in areas such as mental health monitoring, smart surveillance, virtual assistants, and interactive user interfaces. With the rapid advancement of deep learning, significant improvements have been achieved in image-based emotion recognition tasks. Convolutional Neural Networks have proven to be highly effective in extracting

spatial features from images and performing accurate classification. However, one of the key challenges in real-time emotion recognition systems is accurate and efficient face detection before emotion classification. Variations in lighting conditions, facial orientations, occlusions, and background noise make this task more complex in real-world scenarios.

To address these challenges, this paper proposes a robust and efficient facial emotion detection system that combines object detection and deep learning techniques. The system utilizes the YOLO algorithm for fast and accurate face detection, enabling real-time performance. Once the face is detected, a Convolutional Neural Network model is employed to classify the detected face into different emotion categories. The model is trained on a standard facial expression dataset and enhanced through preprocessing techniques such as image resizing, normalization, and data augmentation to improve generalization.

In addition to model development, the proposed system is deployed as an interactive web application to enhance usability and accessibility. The application allows users to perform emotion detection using either static images or real-time webcam input. It also provides confidence scores for predictions and includes an audio feedback feature that verbally communicates the detected emotion, making the system more engaging and user-friendly.

## II. PROBLEM STATEMENT

Accurate recognition of human emotions from facial expressions is a challenging task due to variations in lighting, facial orientation, occlusions, and background noise. Many existing systems either lack efficient real-time face detection or rely on preprocessed images, making them unsuitable for practical applications.

There is a need for a system that can detect faces accurately in real time and classify emotions reliably from both static images and live video. Additionally, improving user interaction through confidence-based outputs and audio feedback is essential for real-world usability.

This project addresses these challenges by integrating a fast face detection approach with a deep learning-based emotion classification model, providing an efficient and user-friendly real-time facial emotion recognition system.

## III. PROPOSED MODEL

The proposed model is designed as a two-stage pipeline that integrates face detection and emotion classification to achieve accurate and real-time facial emotion recognition. The system combines the efficiency of object detection with the learning capability of deep neural networks to process both static images and live video streams.

### A. System Overview

The overall architecture of the system consists of the following major components:

#### 1. Input Module

The system accepts input either in the form of uploaded images or real-time video streams captured through a webcam.

#### 2. Face Detection using YOLO

A YOLO-based object detection model is used to detect and localize faces within the input. This ensures fast and accurate detection even in real-time scenarios and complex backgrounds.

#### 3. Preprocessing

The detected face region is extracted and preprocessed. This includes resizing the image to a fixed dimension, normalization of pixel values, and conversion to grayscale (if applicable) to match the input requirements of the classification model.

#### 4. Emotion Classification using CNN

The preprocessed face image is passed to a Convolutional Neural Network model. The CNN extracts important spatial features and classifies the image into one of the predefined emotion categories.

#### 5. Output Module

The system outputs the predicted emotion along with a confidence score. Additionally, an audio feedback mechanism is used to verbally communicate the detected emotion to the user.

### B. Proposed Model Diagram

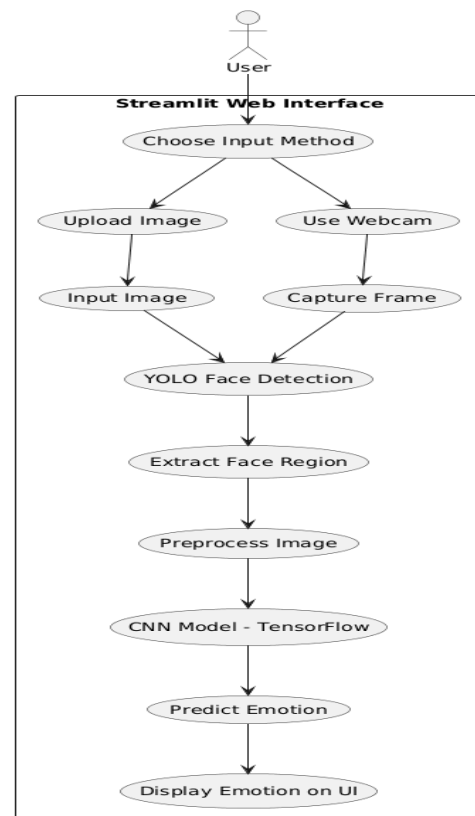


Fig: 3.1: Proposed model

## IV. TECH STACK

The proposed facial emotion recognition system is developed using a combination of modern tools, frameworks, and libraries that support deep learning, computer vision, and web deployment. The selected technologies ensure efficient model development, real-time performance, and user-friendly interaction.

### A. Programming Language

**Python** is used as the primary programming language due to its simplicity, extensive library support, and strong community in machine learning and data science.

## B. Deep Learning Framework

**TensorFlow** is used to build and train the Convolutional Neural Network model for emotion classification. It provides efficient handling of large datasets and supports GPU acceleration.

## C. Computer Vision Library

**OpenCV** is utilized for image processing tasks such as reading images, capturing video through webcam, and handling real-time frame processing.

## D. Face Detection Model

**YOLO (You Only Look Once)** is used for real-time face detection. It enables fast and accurate localization of faces within images and video streams, making the system suitable for real-time applications.

## E. Dataset

The model is trained using the **FER2013 dataset**, which contains labeled facial images categorized into multiple emotion classes.

## F. Web Framework

**Streamlit** is used to develop the interactive web application. It allows easy integration of machine learning models with a user-friendly interface and supports real-time interaction.

## G. Additional Libraries

**NumPy** and **Pandas** are used for data manipulation and preprocessing. **Matplotlib** and **Seaborn** are used for visualization of results and model performance. **Text-to-Speech libraries** are used to generate audio feedback for detected emotions.

## H. Deployment Platform

The application is deployed on the **Streamlit Community Cloud**, making it accessible to users through a web link without requiring local installation.

## V. RESULT SCREENSHOTS

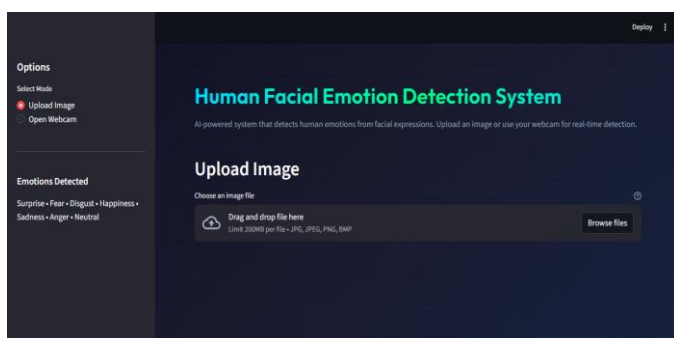


Fig 5.1 : Streamlit webpage

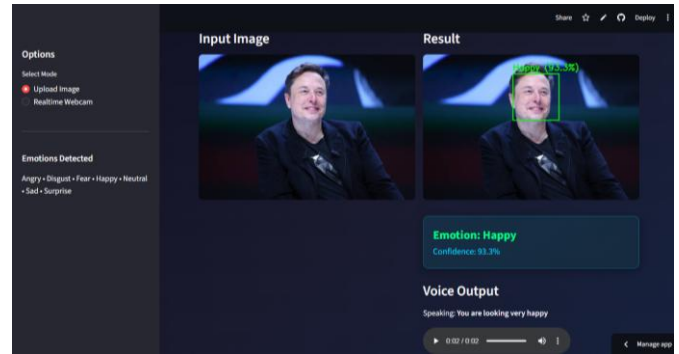


Fig 5.2 : Output

## VI. CONCLUSION

The Human Facial Emotion Detection System developed in this project successfully demonstrates the practical application of deep learning techniques for real-time emotion recognition. The system integrates YOLOv8 for face detection and a Convolutional Neural Network (CNN) for emotion classification, forming a modular and efficient pipeline. The primary objective was to build a system capable of detecting human faces from live webcam input and accurately predicting emotional states in real time. Based on experimental evaluation, the system achieved reliable performance and met the intended goals.

The use of YOLOv8 significantly improved the efficiency of face detection, allowing rapid and accurate localization of facial regions. By isolating only the relevant region of interest (ROI), the system reduces unnecessary background computation and improves overall classification accuracy. The CNN model effectively extracts hierarchical features from facial images and classifies them into predefined emotion categories with satisfactory accuracy.

One of the major strengths of the system is its real-time capability. The application processes frames continuously and displays results instantly, making it suitable for interactive AI-based systems. Additionally, the system operates entirely offline, ensuring privacy and independence from cloud infrastructure.

The major achievements of this project can be summarized as follows:

1. Successful integration of YOLOv8 and CNN in a unified architecture
2. Real-time emotion detection using live webcam feed
3. Modular system design allowing scalability
4. Offline standalone implementation
5. Acceptable classification accuracy with efficient processing speed

The experimental results confirm that deep learning-based emotion recognition systems are capable of delivering reliable and practical solutions in real-world scenarios. While certain environmental factors affect performance, the system demonstrates stable operation under normal conditions.

## REFERENCES

- [1] A. Mollahosseini, D. Chan, and M. H. Mahoor, "AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild," *IEEE Transactions on Affective Computing*, vol. 10, no. 1, pp. 18–31, 2019.
- [2] I. Goodfellow et al., "Challenges in Representation Learning: A Report on Three Machine Learning Contests," *Neural Networks*, vol. 64, pp. 59–63, 2018.
- [3] S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey," *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1195–1215, 2022.
- [4] M. Minaee et al., "Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network," *IEEE Signal Processing Letters*, vol. 27, pp. 1824–1828, 2020.
- [5] Y. Li, J. Zeng, S. Shan, and X. Chen, "Occlusion Aware Facial Expression Recognition Using CNN With Attention Mechanism," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2439–2450, 2019.
- [6] Z. Ge et al., "YOLOX: Exceeding YOLO Series in 2021," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.
- [7] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2018.
- [8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019 edition reference.
- [9] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017 (widely cited foundational CNN paper).
- [10] Ultralytics Documentation, "YOLOv8 Documentation," 2024. Available: <https://docs.ultralytics.com>

