SentifyAI – An AI Emotion Detector

Afifa Shaikh[1], Sufiya Ansari[2], Iqra Essani[3], Noorusabah Sayed[4]

[1,2,3]*Department of Information Technology, M. H. Saboo Siddik Polytechnic, Mumbai, India*
[4]*Head of Department, Department of Artificial Intelligence, M. H. Saboo Siddik Polytechnic, Mumbai, India*

---***---

Abstract - *This paper presents SentifyAI, a web-based system for real-time emotion detection using facial expressions. The application uses OpenCV for face detection and a Convolutional Neural Network (CNN) model to classify emotions from live webcam video. The system detects faces, predicts emotions with confidence scores, and displays the results with emojis in real time. SentifyAI improves human-computer interaction and has potential applications in healthcare, education, and customer service. The system is designed to be simple, fast, and user-friendly for real-time use. It demonstrates the practical use of artificial intelligence in building emotion-aware and interactive systems.*

Key Words: ***Emotion Detection, Deep Learning, Convolutional Neural Network (CNN), OpenCV, Facial Expression Recognition, Human-Computer Interaction, Web Application***.

1. INTRODUCTION

Human emotions are fundamental to communication, decision-making, and social interaction. In traditional computing paradigms, machines operate as passive tools, responding only to explicit commands and remaining oblivious to the user's affective state. This lack of emotional intelligence creates a sterile interaction environment where the machine cannot adapt its responses based on whether the user is frustrated, confused, happy, or engaged. As technology becomes increasingly integrated into daily life in healthcare, education, entertainment, and customer service - the demand for emotionally aware systems has grown significantly.

The field of Affective Computing, pioneered by Rosalind Picard, aims to give machines the ability to recognize, interpret, and simulate human emotions. Recent advancements in Artificial Intelligence (AI), particularly in Deep Learning and Computer Vision, have made it possible to automatically analyze facial expressions, which are a primary channel for expressing emotions. Ekman's foundational work on universal facial expressions identified six basic emotions: happiness, sadness, anger, fear, surprise, and disgust, which serve as the core categories for most recognition systems.

This paper introduces SentifyAI, a web application developed to address the challenge of real-time emotion detection. The system utilizes a standard webcam to capture a video stream. It employs OpenCV for efficient face detection and pre-processing, isolating the region of interest (the face) from each frame. This pre-processed facial image is then passed through a trained Convolutional Neural Network (CNN) model, which classifies the expression into one of the predefined emotion categories.

The detected emotion and its associated confidence score are then displayed back to the user on the video feed, often accompanied by a representative emoji for enhanced user experience. This project demonstrates a complete pipeline from image acquisition to real-time affective feedback, showcasing the potential of integrating deep learning models into accessible web-based platforms.

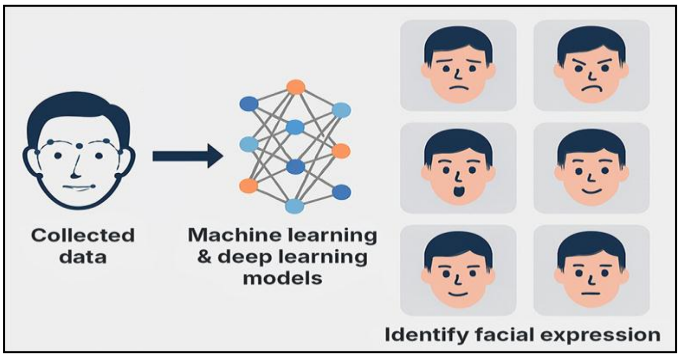

Fig -1: Working of SentifyAI

2. PROBLEM STATEMENT

Traditional computer systems and interfaces are unable to effectively perceive and interpret human emotions. This limitation reduces the quality of human-computer interaction and makes communication less natural and less personalized.

The absence of emotional awareness creates a gap between humans and machines, resulting in communication that lacks empathy and responsiveness. As a result, users may feel less engaged and understood when interacting with digital systems.

In important applications such as teletherapy, online education, and automated customer support, the inability to recognize a user's emotional state can lead to misunderstandings and reduced service quality. This can ultimately result in lower user satisfaction.

Therefore, there is a strong need for an intelligent automated system that can detect and analyze human emotions through facial expressions in real time. Such systems can help create more adaptive, interactive, and user-friendly digital environments.

3. LITERATURE SURVEY

The domain of emotion recognition has evolved significantly over the past decades, transitioning from psychological studies to complex computational models.

a. Psychological Foundations: The groundwork for automatic emotion recognition was laid by psychologists like Paul Ekman, who conducted cross-cultural studies and established that certain basic emotions (happiness, sadness, anger, fear, surprise, disgust) are universally expressed and recognized through specific facial configurations, known as Action Units in the Facial Action Coding System (FACS).

b.Traditional Machine Learning Approaches: Early computational approaches relied on handcrafted feature extraction techniques such as Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), or Gabor filters. These features were then fed into classifiers like Support Vector Machines (SVM) or Random Forests. While effective to a degree, these methods were sensitive to variations in lighting, pose, and occlusion and required careful feature engineering.

c. Deep Learning Revolution: The advent of Deep Learning, particularly Convolutional Neural Networks (CNNs), revolutionized the field. CNNs can automatically learn hierarchical features directly from raw pixel data, eliminating the need for manual feature extraction. Architectures like VGGNet, ResNet, and custom smaller models have been successfully applied to facial expression recognition datasets such as FER-2013, CK+, and AffectNet, achieving state-of-the-art accuracy.

d. Existing Systems: Current commercial systems like Microsoft Azure Face API and Amazon Recognition offer emotion detection as part of broader facial analysis services. However, they are often cloud-based, introducing latency and privacy concerns. SentifyAI differentiates itself by aiming for a lightweight, real-time, and potentially privacy-preserving (if run locally) web broader facial analysis services. However, they are often cloud-based, introducing latency and privacy concerns.

4. METHODOLOGY

The development of SentifyAI followed a systematic pipeline, integrating computer vision and deep learning for real-time performance.

a. Data Acquisition and Preprocessing: A crucial step for any deep learning model is the data. The system utilizes a publicly available dataset for training, such as FER-2013 (Facial Expression Recognition 2013), which contains grayscale images of faces labelled with seven emotion categories (Angry, Disgust, Fear, Happy, Sad, Surprise, Neutral). Preprocessing involves converting frames from the webcam to grayscale, detecting the face using OpenCV's Haar Cascade classifier, cropping the face region, and resizing it to the dimensions expected by the CNN model (e.g., 48x48 pixels).

b. Model Architecture (CNN): The core of SentifyAI is a Convolutional Neural Network. The architecture consists of:

Convolutional Layers: Multiple layers with small filters (e.g., 3x3) to extract features like edges, textures, and complex facial patterns. ReLU (Rectified Linear Unit) activation is used to introduce non-linearity.

Pooling Layers: Max-pooling layers follow convolutional layers to reduce the spatial dimensions, decrease computational load, and make feature detection more robust.

Fully Connected Layers: After several convolutional and pooling layers, the feature maps are flattened and passed through one or more dense layers to perform high-level reasoning.

Output Layer: A final dense layer with 7 neurons (for 7 emotions) and a SoftMax activation function, which outputs a probability distribution over the emotion classes.

c. Model Training: The CNN model is trained on the pre-processed dataset using a categorical cross-entropy loss function and an optimizer like Adam. The goal is to minimize loss and maximize classification accuracy on a validation set.

d. Real-Time Inference and Integration: The trained model is saved and loaded into the main application. The application captures video frames from the webcam using OpenCV. For each frame, face detection is performed. For every detected face, the region is pre-processed and fed into the CNN model. The model predicts the emotion and its confidence score. OpenCV is used to draw a bounding box around the face and display the emotion label, confidence score, and a corresponding emoji on the frame in real-time.

5. SYSTEM ARCHITECTURE

The architecture of SentifyAI follows a modular pipeline, processing data sequentially from input to output. Input Module (Webcam) captures real-time video feed. Face Detection Module (OpenCV) employs a pre-trained Haar Cascade classifier to locate faces within each frame. The face region is cropped. Preprocessing Module converts the cropped face to grayscale (if not already), resizes it to the target dimensions (e.g., 48x48), and normalizes pixel values. Emotion Classification Module (Trained CNN Model) takes the preprocessed face image as input and outputs a probability for each emotion class. The class with the highest probability is selected. Visualization Module (OpenCV) draws the results (bounding box, emotion label, confidence, emoji) onto the original video frame. Display Module shows the processed video stream to the user via the web application interface.

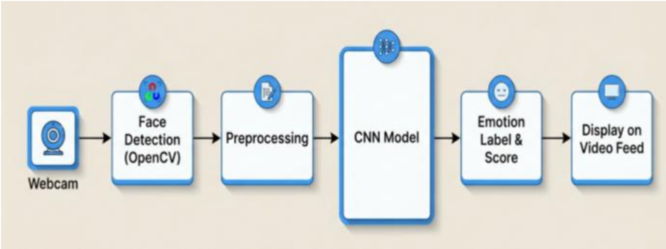

Fig – 2: System Architecture of SentifyAI

6. SYSTEM REQUIREMENTS

Software Requirements:

Operating System: Windows, Linux, or macOS
Programming Language: Python 3.x
Libraries: OpenCV, TensorFlow/Keras,Flask, NumPy
Frontend: HTML, CSS, JavaScript
Development Environment: VS Code / PyCharm / Jupyter Notebook

Hardware Requirements:

Processor: Intel Core i5 / AMD Ryzen 5 or higher (for smooth real-time processing)
RAM: Minimum 8 GB (16 GB recommended)
Camera: Standard USB Webcam (720p or higher)
GPU (Recommended): NVIDIA GPU with CUDA support to accelerate model inference
Storage: Minimum 2 GB free space for code and libraries

7. OUTPUT & RESULTS

The primary output of the system is a real-time, annotated video stream. The expected outputs include:

Emotion Label: Text overlay displaying the predicted emotion (e.g., "Happy", "Sad"). Confidence Score: A percentage value indicating the model's certainty in its prediction. Visualization: A representative emoji corresponding to the detected emotion displayed next to the label for intuitive understanding. Bounding Box: A rectangle drawn around the detected face.

8. CONCLUSION

SentifyAI successfully demonstrates the feasibility of a real-time, web-based emotion detection system using accessible technologies like OpenCV and CNNs. The project achieves its primary objective of bridging the gap in human-computer interaction by enabling machines to perceive and display human affective states. This system serves as a foundational model for integrating emotional intelligence into various applications.

REFERENCES

[1] M. M. Hassan, G. Muhammad, and M. S. Hossain, “Affective Computing in Human–Computer Interaction: A Review,” **IEEE Access**, vol. 7, pp. 117393–117406, 2019.

[2] S. Li and W. Deng, “Deep Facial Expression Recognition: A Survey,” **IEEE Trans. Affective Comput.**, vol. 13, no. 3, pp. 1195–1215, 2022. https://doi.org/10.1109/TAFFC.2020.2981446

[3] S. Wang, Z. Liu, Y. Lv, Y. Zhu, and F. Wu, “A Survey on Facial Expression Recognition: From Traditional to Deep Learning Methods,” **IEEE Trans. Affective Comput.**, vol. 10, no. 3, pp. 394–409, 2019.

[4] Mollahosseini, B. Hasani, and M. H. Mahoor, “AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild,” **IEEE Trans. Affective Comput.**, vol. 10, no. 1, pp. 18–31, 2019. https://doi.org/10.1109/TAFFC.2017.2740923

[5] E. Barsoum, C. Zhang, C. C. Ferrer, and Z. Zhang, “Training Deep Networks for Facial Expression Recognition with Crowd-Sourced Label Distribution,” Proc. ICMI, 2016.

[6] J. Chen, Z. Chen, Z. Chi, and H. Fu, “Facial Expression Recognition in the Wild Using Improved CNN,” **IEEE Access**, vol. 8, pp. 190652–190660, 2020.

[7] M. Minaee, M. Minaei, and A. Abdolrashidi, “Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network,” **Sensors**, vol. 21, no. 9, 2021. https://doi.org/10.3390/s21093116

[8] S. S. S. R. Kumar and P. V. Arun, “Real-Time Facial Emotion Recognition Using Deep Learning,” **Electronics**, vol. 8, no. 10, 2019. https://doi.org/10.3390/electronics8101178

[9] X. Li, J. Chen, D. Zhao, and J. Du, “Lightweight Deep Learning Model for Real-Time Facial Expression Recognition,” **Applied Sciences**, vol. 11, no. 21, 2021. https://doi.org/10.3390/app11219278