

# Neuro Bio Mood: AI-Driven Emotion Analysis from Biometric Cues

<sup>1</sup>Dr. K. Premkumar, <sup>2</sup>Maheswaran S, <sup>3</sup>Prathik Sarathy V\*, <sup>4</sup>Dharunprakash A, <sup>5</sup>Sri Avinash T

<sup>1</sup>Head of the Department, Department of Computer Science and Engineering, Sri Manakula Vinayagar Engineering College, Madagadipet, Puducherry, India 605-017

<sup>2,3,4,5</sup> UG Student, Department Of Computer Science and Engineering, Sri Manakula Vinayagar Engineering College, Madagadipet, Puducherry, India 605-017

DOI: <https://doi.org/10.51584/IJRIAS.2025.10040016>

Received: 11 April 2025; Accepted: 22 April 2025; Published: 29 April 2025

## ABSTRACT

This project is about "Emotion Detection using Human Biometric Inputs through Machine Learning" which aims in human emotion classification or recognition using various human biometric and physiological signals. Combination of deep learning and traditional machine learning techniques to extract feature for emotion classification, which helps to improve overall recognition accuracy. Deep learning models are used to decode facial expressions, voice intonations, and natural language semantics and machine learning algorithms assist with accurate classification of emotional signals. It, therefore, creates a more wholesome understanding of human emotion, making the whole system more adaptable and reliable in nature. Applications could include health care (monitoring emotional response), human-computer interaction (designing systems that respond more appropriately to emotional state), and possibly security (detecting stress or suspicion) Viza Emotion thus serves to posses the foundations of more precise, intelligent, integrating a heuristic and emotional approach to technological solutions.

**Index Terms:** Emotion Detection, Human Facial Inputs, Human Speech Input, Human input text, Deep Learning. Conventional Machine Learning.

## INTRODUCTION

In recent years, the capacity of machines to understand human emotions and respond appropriately has been an essential part of creating intelligent and interactive systems. Emotion detection is a meta stage of Affective Computing, it is the study of human activity and human emotion from data infused with physiological, behavioral and contextual signal. Facial recognition and voice recognition are both well-known for emotion analysis as the two most natural and non-invasive modalities among various approaches.

Facial expressions are a shared language of emotion, able to communicate everything from happiness to anger, fear to sadness. In a like manner, the human voice reflects subtle emotional states that may not always manifest in the face, be it through modulations of pitch, tone, and rhythm. These two powerful cues used together provide a more complete and accurate version of human emotions.

In this project, we design and build a machine learning powered face and human voice recognition based emotion detection system. This system is trained on datasets, where facial images and its speech samples are paired with respective labels. By combining both modalities, the detection process becomes more efficient and accurate, which can be useful in many applications such as mental health monitoring, virtual assistants, e-learning platforms, and human-computer interaction.

In this research, we are going to investigate the multi-modal emotion recognition issues, algorithms, and efficacy by applying the state of the art machine learning methods. We work towards this aim in a hope to develop emotionally intelligent systems capable of responsive interaction recognising human needs.

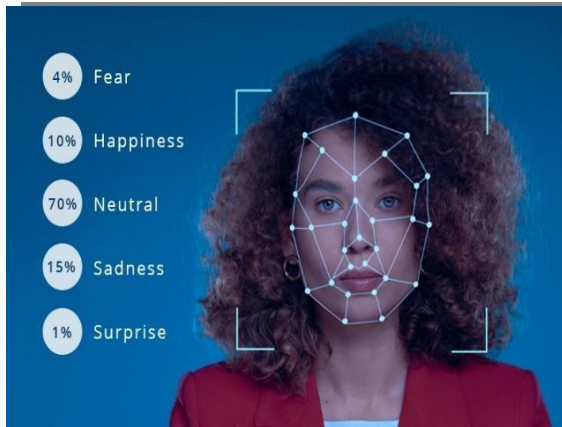


Fig 1.1 Classification of Human Emotions

### Challenges for the Emotion Detection Systems:

There are multiple hurdles emotion detection systems cross, mainly arising from diversity in data and environmental intermediates. Because people express emotions so differently across a range of cultural, personal, and situational factors (what is a sign of joy in one culture is just being polite in another) it is hard to generalise any machine learning models when there is many types of human population. Recognition systems are affected negatively by environmental conditions such as low lighting, noise in the background, closed circuits or occlusions about facial. To extract facial cues use machine learning and convolutional neural nets (CNN) to do facial expression recognition and for voice emotion detection interpret the vocal patterns, tone & pitch which could be distorted with individual voice characteristics or external noise. Moreover, the machine learning-based models typically work as "black boxes" which makes difficult to dissect the decision process and interpretability of AI-driven systems has been one of the greatest concerns.

### Proposed Systems:

They suggest a multi-modal emotion detection system that involves reading facial and voice expressions for the accurate detection of human feelings. The system allows a simultaneous feed of video from the user's webcam and his voice from the microphone, then processes the former through a Convolutional Neural Network (CNN) and the latter by applying the Mel-Frequency Cepstral Coefficients (MFCCs), pitch, and tone + the classifiers such as Support Vector Machines (SVM) or Recurrent Neural Networks (RNNs). The responses received from both modes are fused based on a late fusion scheme, which aggregates decisions at the level of the final prediction to increase precision and dependability. The interface not only has a simple, easy-to-understand graphic display of user emotions but actually is a tool that makes real-time human-computer interaction, virtual assistants, online education, and mental health monitoring possible.

### Benefits of Emotion Detection in the Medical Field:

Advanced emotion detection technologies have made a gradual revolution in patient care, mental health evaluation, and treatment approaches in the medical field. These AI-powered tools help the chemists and the druggists diagnose and treat mental health conditions like anxiety and depression by using biometric sensors and deep learning algorithms to stay in track of the human emotional states in real time. Emotion detection aids in supervising the patients' stress and pain levels in critical care units, allowing for prompt interventions. Advanced emotion detection systems enhance the medical decision-making, individualized patient care, and overall healthcare efficiency by making use of the AI and real-time data processing.

**a) Mental Health Monitoring:** Early intervention through the analysis of emotional disturbance.

**b) Human-Computer Interaction:** Enhancing the games and virtual helpers.

**c) Customer Sentiment Analysis:** Improving the market research and customer support.

**d) Education:** Customizing the educational experiences in accordance to the feelings of students.

## Problem Definition

Human communication relies heavily on the human emotions and also feelings, even though it is still difficult to reliably identify emotions from many physiological inputs. The main aim of this project is to develop an emotion detection system that can efficiently identify emotions by analyzing text input, facial expressions, and human speech. The system can also identify the human emotions including happy, sorrow, rage, and neutrality by combining audio processing, facial recognition, and natural language processing (NLP). By offering a more user-friendly and responsive AI-driven emotion recognition system, the objective is to enhance the sentiment-based applications, mental health analysis, and human-computer interaction.

## LITERATURE SURVEY

Swadha Gupta, Raj Kumar Tekchandani, and Parteek Kumar suggested a notion of designing a real-time learner engagement system for the context that had some disadvantages, such as the emotion detection system could identify some basic emotions like Happy, Sad, and Anger but did not use the relevant cues like the voice tone, body language, or eye movement. To mitigate that issue, we added some more nuanced emotion classes to engagement such as "Focused," "Distracted," "Bored," and "Confused" and would use methodologies, such as Federated Learning, Differential Privacy, and Face Blurring options to protect the identity of the user.

Amjath Rehman Khan proposed a framework for Facial Recognition via Conventional Machine Learning and Deep Learning methods that had limitations whereby the model had troubles detecting and classifying micro-expressions from human face, which in some cases lasted for milliseconds, in addition to the computational costs of using Atrous CNN's and RNN's. In addition, to reduce the computation time and costs to classify micro-expressions in real-time, we supplemented the training data by using generative adversarial networks to produce additional micro-expressions to increase and balance the training data. Further, we utilized transfer learning methods to pre-train and fine-tune models using existing large data facial emotion datasets or few-shots learning models.

Zhenjie Song also suggested an idea of the Facial Expression Emotion Recognition Model that is a consequence of the combination of philosophy and theory of machine learning. However, the system that was based on the FER2013 database had drawbacks including low resolution, noisy labels, and a lack of variation in facial features across demographics such as ethnicity, age, and lighting conditions. Additionally, the Gabor filters that are applied for feature extraction created an additional pre-processing overhead and do not automatically scale well for real-time systems or deployment on the edge devices. Therefore, we used varied data sets such as AffectNet, RAF-DB, or the EmotioNet in order to build robustness into the data model and improve adaptability to the real-world environment. We also created an ethical reasoning layer or rule-based interpretation interface to realize the philosophical values and facilitate fairness in the data models and trust in the process and ultimate outcomes.

Fathi E. E. Abd El-Samie and Nagla F. Soliman introduced a concept of using machine learning techniques for human emotion detection. They acknowledged the disadvantages such as KNN algorithm having a slow inference time and poorly scaling with larger datasets, which made it impractical for real-time robotic vision in real-life dynamic situations and referred to the segmented processing stage of human emotion detection, which limited the model to learn the optimal expressions holistically unlike end-to-end deep learning models. Therefore, in order to improve this, we replaced this KNN algorithm with Support Vector Machines, to improve speed and scalability, along with using Recurrent Neural Networks and Temporal Convolutional Networks to account for emotional transitions over time, especially used for the robots that have continuous interaction.

Priti Verma, Mejdal Alqahtani, Prathibha S, and Awal Halifa proposed an approach that employed the facial emotion recognition techniques in the real-world scenarios, where there were disadvantages like the model in the system being heavily reliant on the standard datasets, e.g. the ImageNet, which are general datasets for the learning paradigm that do not focus on the emotions; the model could have some dismorphism when translating to the emotions these images; and the model might not generalize across the different ages, ethnicities, and facial structures unless the training and augmentation was sufficient - ultimately leading to incorrect bias or poor

performance in their study. So in order to mitigate this, the authors integrated the self-attention and the transformer-based modules to more accurately focus on the relevant facial regions in order to improve the robustness of the model towards occlusions and distortions, along with fairness-aware learning to consider equity in model performance across the different demographics by balancing the datasets, and evaluating with bias-detection measurements.

The title of the research by Monisha G S, Yogashree, Baghyalakshmi R, and Haritha P is "Enhanced Automatic Recognition of Human Emotions Using Machine Learning Techniques". As a major setback, the model trained using AffectNet has achieved validation accuracy only up to 56.54%, which indicates that it is very difficult to generalize across different datasets. Hence, domain adaptation techniques have been adopted with significantly diverse datasets to enable better generalization by the model.

Kristina Machova, Martina Szaboova, Jan Paralic, and Jan Micko have put forward the Detection of Emotion by Text-Analysis using Machine Learning, which has demerits such as the nonliteral interpretations by the model that lead to misclassification of emotions, such as sarcasm, irony, idioms, and metaphors. Then we resort to advanced natural language processing methods, such as transformer-based models (e.g., BERT, GPT), which work better in capturing underlying contextual nuances and interpreting figurative language.

### Architectural Diagram:

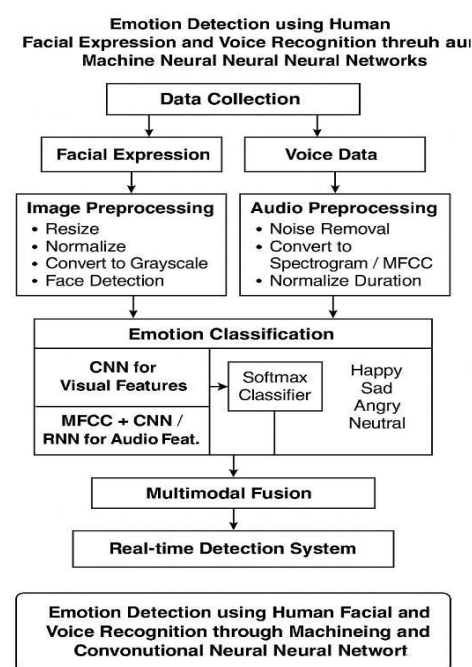


Fig 1.2 Architectural Diagram

### Process of Emotion Detection via the Human Facial Expression and Voice Recognition:

The list of processes involved in the Emotion Detection via Human Face and Voice Recognition are listed below as follows:

#### (1) Data collection:

Data collection involves gathering labeled datasets that include many human emotions (happy, sad, angry, neutral). For facial expression recognition, several datasets are available such as FER-2013, AffectNet, CK, and JAFFE, which contain images of faces marked with the appropriate emotions. In the case of detecting emotion based on voice, datasets such as RAVDESS, SAVEE, and CREMA- D can be used to supply audio samples with annotations of emotions to allow for training and evaluation of machine learning models.

**(2) Data Pre-processing:** The goal of data preprocessing is to clean the data and prepare the data in a format necessary for training a model. For facial expression data, preprocessing consists of reducing the dimensions of

the images to a standard dimension such as 48x48 or 224x224, normalizing pixel values found in an image, converting an image to grayscale (if needed) and performing face region detection from a static video frame or image (by utilizing Haar Cascades or Dlib). For voice data specifically it consists of a few steps which are removing background noise from a data sample, converting an audio recording into a visual representation with spectrogram, MFCC (Mel Frequency Cepstra Coefficients) or Mel specifications and normalizing the duration sample to provide consistency across data samples.

### **(3) Feature Extraction:**

The purpose of Feature Extraction is to gather patterns from image and audio data that hold meaning and are discriminative for emotion recognition. In the case of facial expression-based analysis, Convolutional Neural Networks (CNNs) will be used to learn visual features from facial images automatically. For voice emotion detection, we will extract audio features known as Mel Frequency Cepstral Coefficients (MFCCs), and then these features will be analyzed using CNNs or 1D CNN and Long Short Term Memory (LSTM) networks to capture spatial and temporal information from the emotional speech signals.

### **(4) Emotion Classification:**

Emotion classification is the training of a machine learning model to successfully classify emotions in the extracted features from facial expressions and voice. Facial expressions can be recognized using deep CNN architectures with Softmax layers at the output, examples of deep CNN architectures include VGG16, ResNet, Inception-ResNet, and MobileNet. In detecting emotions based on voices, MFCC features can be extracted and combined with CNN, LSTM, or Bi-LSTM architectures, capturing temporal and spectral patterns. Classifications can then be done using Softmax or a traditional machine learning classifier (Support Vector Machine-SVM).

### **(5) Real-Time Detection System:**

The Real-time Detection System is concerned with implementing the previously trained emotion recognition models into an active system for real-time use with a camera and microphone. OpenCV and similar tools will be used to capture facial expressions through a camera, while SpeechRecognition, PyDub, or Librosa will be used to input voice in real time. To minimize latency, lighter weight versions of any previously trained models will be used for inference, and the recognized emotions will be displayed on the screen in real time.

### **(6) Evaluation & Optimisation:**

The Evaluation & Optimization phase will focus on measuring the performance and accuracy of the emotion detection models on real-world data. Models will again be evaluated using metrics, including accuracy, precision, recall, F1-score, confusion matrix, and ROC-AUC curve, and then optimize the model to further improve performance and robustness. Model optimization can include switching, tuning hyperparameters, and/or pruning or quantizing the model, particularly to improve real-time deployment efficiencies.

### **Evaluation Metrics in the Emotion Detection System:**

In order to evaluate the Emotion Detection System based on the Human Facial Expression and the Voice Recognition Recognition, several Evaluation Metrics are deployed to accurately measure its effectiveness without any deviations. Those evaluation metrics are:

#### **(a) Precision:**

It evaluates the model's performance in correctly classifying the positive classes and penalizes it for mistakes in classifying negative classes.

#### **(b) Recall:**

It evaluates the model's ability to detect all actual occurrences of emotion.



### (c) F1-score:

It indicates a balanced measure of precision and recall when the possible outcomes are imbalanced across classes since it is the harmonic mean of both precision and recall.

### (d) Confusion Matrix:

It represents the performances across all classes of emotion, including both hits and misses.

### (e) Receiver Operating Characteristic – Area Under Curve:

It tests how well the model differentiates between classes in button or multimodal cases.

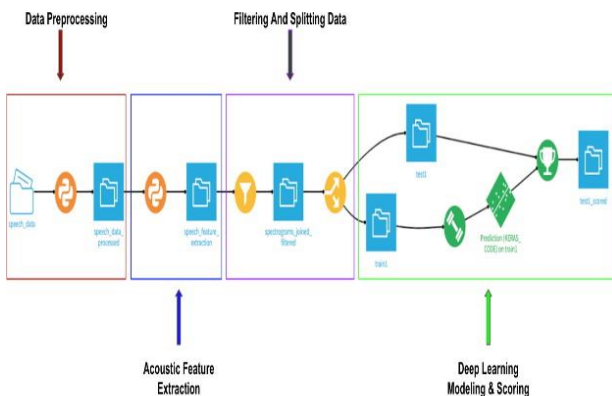


Fig 1.3 Process Workflow of Emotion Detection using Human Voice Recognition

### Comparison Diagram of Our Project with the Existing Project:

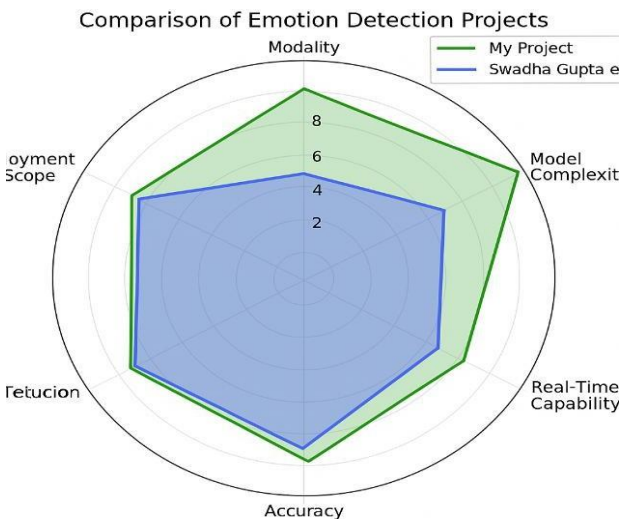


Fig 1.4 Comparison of our project with a existing project

### Project Implementation

The project begins by gathering and preparing datasets for textual emotions (ISEAR, Sentiment140), facial expressions (FER2013, CK+), and voice data (RAVDESS, EmoDB). The data is preprocessed by performing tokenization and embedding for text, MFCC feature extraction for audio, and grayscale conversion and resizing for facial images. Features are then extracted using **Convolutional Neural Networks (CNN)** for facial expressions, **Recurrent Neural Networks (RNN)** or LSTM for speech emotion recognition, and Natural Language Processing methods like Word2Vec and BERT for text sentiment analysis. Using deep learning frameworks such as **TensorFlow** and **PyTorch**, distinct models are trained for each biometric input, and Multi-Task Learning (MTL) is employed to integrate knowledge from all modalities. For emotion classification, softmax



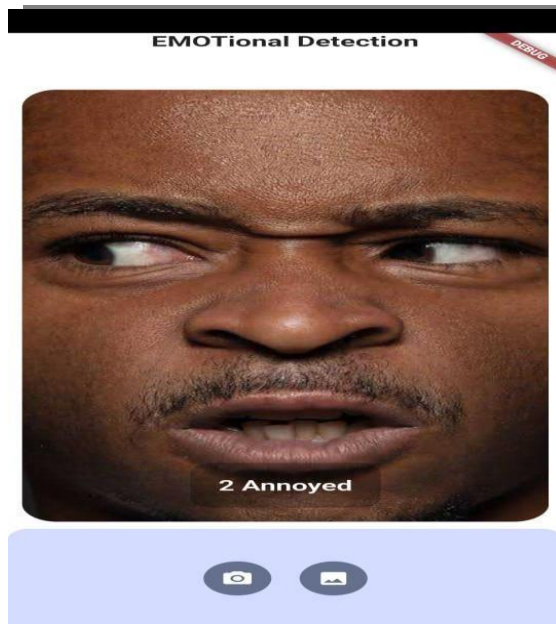


Fig 1.7 Output for converting the text of any language into English and also recognising the emotions in the human speech input

## CONCLUSION

Emotion recognition through human biometrics using machine learning has revolutionized the way artificial intelligence interprets human emotions. It hence enables accurate as well as real-time emotion detection based on an enriched array of biometric data, which includes facial expressions, voice tone, heart rate, EEG signals, and other physiological indicators. The performance both in terms of accuracy and flexibility of the system to detect emotions has been significantly improved by employing deep learning models such as CNNs, RNNs, and transformers.

The future of emotion detection technology will rely on all these combinations of advances in AI, neuroscience, quantum computing, and nanotechnology that will eventually permit more robust, real-time adaptive context-sensitive emotion recognition systems. The BCIs, smart environments with affective computing capabilities, and AI-driven therapies for emotional regulation are all permitted by the technology called to have tremendous implications. Continued development as a discipline will continue to play an important role in clarifying the confusion over human emotions versus AI to make interactions deeper and more empathetic across sectors. Of course, ethical questions in addressing data privacy and protection as well as the fair functioning of AI models will remain critical issues for the future.

## REFERENCES

1. "Facial Emotion Based Real-Time Learner Engagement System in Online Learning Context using the Deep Learning Models" |Swadha Gupta, Parteek Kumar, Raj Kumar Tekchandani | 2022 | doi: 10.1007/s11042-022-13558-9
2. "Facial Emotion Recognition Using Conventional Machine Learning and Deep Learning Methods: Current Achievements, Analysis and Remaining Challenges"| Amjad Rehman Khan | 2022| doi: 10.3390/info13060268
3. "Deploying Machine Learning Techniques for Human Emotion Detection" | Ali I. Siam, Naglaa F. Soliman, Abeer D. Algarni, Fathi E. Abd El-Samie, Ahmed Sedik| 2022| doi: 10.1155/2022/8032673
4. "Methods for Facial Expression Recognition with Applications in Challenging Situations" |Anil Audumbar Pise, Mejdal A. Alqahtani, Priti Verma, Purushothama K, Dimitrios A. Karras, Prathibha S, Awal Halifa| 2022| doi: 10.1155/2022/9261438
5. "Facial Expression Emotion Recognition Model Integrating Philosophy and Machine Learning Theory" |ZhenjieSong|2021|doi: 10.3389/fpsyg.2021.759485
6. "A study on computer vision for facial emotion recognition" |Zi-Yu Huang, Chia-Chin Chiang, Jian-Hao



- 
- Chen, Yi-Chian Chen, Hsin-Lung Chung, Yu-Ping Cai, Hsiu-Chuan Hsu | 2023| doi: 10.1038/s41598-023-35446-4 “Enhanced Automatic Recognition of Human Emotions Using Machine Learning Techniques”
7. |Monisha.G.S, Yogashreee.G.S, Baghyalakshmi.R, Haritha.P | 2023| doi: 10.1016/j.procs.2023.01.020
  8. “Detection of Emotion by Text Analysis using Machine Learning” |Kristina Machova, Martina Szaboova, Jan Paralic, Jan Micko | 2023| doi: 10.3389/fpsyg.2023.1190326 “Emotion Detection using Machine Learning”
  9. |Ashadu Jaman Shawon, Anika Tabassum, Rifath Mahmud | 2023| doi: 10.56532/mjsat.v4i1.195
  10. “A review on Emotion Detection by using the Deep Learning Techniques” |Tulika Chutia, Nomi Baruah | 2024| doi: 10.1007/s10462-024-10831-1
  11. “Machine Learning Techniques for Emotion Detection and Sentiment Analysis: Current State, Challenges, and future directions” |Alaa Alslaity, Rita Orji|2022|doi: 10.1080/0144929X.2022.2156387
  12. Facial Emotion Detection using Machine Learning” |Prof.Sopan Kshirsagar, Harshad Shinde, Salman Shikalgar, Ruturaj Raut |2024| doi: 10.1038/s41598-023-35446-4