International Journal for Modern Trends in Science and Technology
Volume 12, Issue 04, pages 902-910
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue04.html
DOI: https://doi.org/10.5281/zenodo.19644334

Emotion Aware Human Computer Interaction Using Real-Time Facial Expression Analysis

L. Dana Ratna Kishore, T. Roopa, V. Geethanjali, Y. S. D. N. Ganesh, M. Sai Hemanth Varma

Department of Computer Science and Engineering, D.N.R. College of Engineering & Technology, Balusumudi, Bhimavaram, Andhra Pradesh, India

To Cite this Article
Mr. L. Dana Ratna Kishore, T. Roopa, V. Geethanjali, Y. S. D. N. Ganesh & M. Sai Hemanth Varma (2026). Emotion Aware Human Computer Interaction Using Real-Time Facial Expression Analysis. International Journal for Modern Trends in Science and Technology, 12(04), 9022-910. https://doi.org/10.5281/zenodo.19644334

Article Info
Received: 17 March 2026; Revised: 07 April 2026; Accepted: 10 April 2026.

KEYWORDS

Facial Emotion Recognition (FER), Convolutional Neural Network (CNN), Pleasure–Arousal Dominance (PAD) model, Real-time Emotion Detection, Python, TensorFlow, OpenCV, Deep Learning, Human-Computer Interaction, Emotion Classification

ABSTRACT

Facial Emotion Recognition (FER) helps computers understand human emotions, but existing systems rely on traditional machine learning methods that require manual feature extraction and provide lower accuracy (around 60–70%), making them unsuitable for real-time use. To overcome this problem, this project proposes a real-time FER system using a Convolutional Neural Network (CNN) that automatically extracts facial features and maps them to the Pleasure–Arousal–Dominance (PAD) model. The system is implemented using Python with TensorFlow, OpenCV, NumPy, Pandas, and Streamlit. Experimental results show 94% accuracy for the Pleasure dimension and around 73% accuracy for overall emotion classification. The system works in real-time and can be used in healthcare, education, and smart monitoring applications.

1. INTRODUCTION

The rapid advancement of Artificial Intelligence (AI) and Human-Computer Interaction (HCI) technologies has significantly transformed the way humans interact with machines. Traditional computer systems are primarily designed to respond to explicit user commands such as text, voice, or predefined inputs. However, these systems lack the ability to understand human emotions, which are a fundamental aspect of natural communication. This limitation reduces the effectiveness and intuitiveness of human-computer interaction [1].

To address this challenge, emotion-aware systems have emerged as an important research area. Emotion recognition enables machines to detect and interpret human emotional states, allowing systems to respond more intelligently and naturally. Among various approaches, facial expression analysis is considered one of the most effective and non-intrusive methods for emotion detection, as human faces convey rich emotional information [2].This project, titled *"Emotion Aware Human Computer Interaction using Real-Time Facial Expression Analysis,"* focuses on developing an intelligent

system capable of detecting human emotions in real time. A webcam is used to capture live video input, from which human faces are detected using image processing techniques. Key facial features such as eyes, eyebrows, and mouth are extracted and analyzed to identify emotional cues. These features are then processed using a Convolutional Neural Network (CNN), a deep learning model trained to classify emotions such as happy, sad, angry, surprised, and neutral [3].The integration of real-time processing with deep learning techniques enables the system to achieve high accuracy and responsiveness. This allows machines to adapt dynamically based on user emotions, making interactions more natural and user-friendly. Emotion-aware systems have wide applications in domains such as smart classrooms, healthcare monitoring, driver safety systems, virtual assistants, and interactive gaming, where understanding user emotions enhances performance and user experience [4].

Emotion recognition has evolved from basic rule-based systems to advanced deep learning approaches. Earlier systems relied on manual feature extraction methods, which limited their accuracy and scalability. In contrast, modern CNN-based models automatically learn hierarchical features from images, significantly improving performance and enabling real-time applications [5]. As a result, emotion-aware systems play a crucial role in bridging the gap between human feelings and computer responses, leading to more intelligent and empathetic interactions.

Purpose

The purpose of this project is to design and develop a real-time emotion detection system that enhances human-computer interaction by enabling machines to understand human emotions through facial expressions.

The proposed system aims to:

- Capture live facial data using a webcam for real-time processing
- Detect and extract facial features using image processing techniques
- Classify emotions such as happy, sad, angry, surprised, and neutral using a CNN model
- Enable intelligent system responses based on detected emotional states

By leveraging deep learning and real-time processing, the system improves interaction quality, making it more natural, adaptive, and user-friendly. This approach is particularly useful in applications such as education, healthcare, and smart systems, where understanding user emotions can significantly enhance service quality and user experience [3], [4].

Motivation

In the modern digital era, computers and intelligent systems are capable of performing complex tasks with high efficiency. However, most existing systems still lack the ability to understand human emotions, which creates a significant gap in human-computer interaction.

Humans naturally express emotions through facial expressions, and the ability to interpret these emotions is essential for meaningful communication. Without emotional awareness, computer systems remain limited in providing personalized and context-aware responses. This lack of emotional intelligence reduces user satisfaction and limits the effectiveness of interactive systems [2].The motivation behind this project is to bridge this gap by developing a system that can recognize human emotions in real time using facial expressions. By understanding the emotional state of users, systems can adapt their behavior, provide personalized feedback, and improve overall interaction quality. This leads to more natural, efficient, and engaging user experiences across various applications [4].

Problem Statement

Traditional computer systems are designed to process explicit inputs and lack the capability to interpret human emotions, resulting in limited interaction quality. Although some emotion recognition systems have been developed, many of them face significant challenges such as lack of real-time performance, low accuracy, and dependence on manual feature extraction techniques.

As a result:

- Systems cannot effectively understand user emotions during interaction
- Many existing models fail to perform accurately in real-time environments
- Traditional approaches rely on handcrafted features, limiting scalability and robustness
- User experience remains less interactive and less personalized

Therefore, there is a need for an intelligent system that can detect and classify human emotions accurately in real time using automated feature learning techniques. The proposed CNN-based emotion recognition system

aims to overcome these limitations by providing improved accuracy, adaptability, and responsiveness, ultimately enhancing human-computer interaction and user experience [5].

2 Literature Survey

[1] ***Facial Expression Recognition using Machine Learning***

Paul Ekman and Wallace V. Friesen introduced one of the earliest approaches for emotion recognition based on Facial Action Coding System (FACS). Their work focuses on identifying human emotions by analyzing facial muscle movements known as action units. Traditional machine learning algorithms such as Support Vector Machines (SVM) were used for classification. Although effective for controlled environments, this approach relies heavily on manual feature extraction and struggles with variations in lighting, pose, and real-time implementation.

[2] ***Real-Time Emotion Detection using Convolutional Neural Networks (CNN)***

Ian Goodfellow and Yoshua Bengio explored the application of deep learning techniques, particularly Convolutional Neural Networks (CNN), for emotion detection. CNN models automatically extract relevant features from facial images, eliminating the need for manual feature engineering. This approach significantly improves accuracy and enables real-time emotion classification using webcam inputs. However, the model requires large datasets and high computational resources for training and deployment.

[3] ***Emotion Recognition from Facial Expressions using Deep Learning***

Ali Mollahosseini and David Chan proposed a deep learning-based framework for facial emotion recognition using large-scale datasets. Their model demonstrates improved accuracy in detecting multiple emotions under varying conditions such as different lighting and face orientations. The approach enhances robustness compared to traditional methods, but real-time performance may be affected when processing high-resolution video streams.

[4] ***Real-Time Face Detection using Haar Cascade Algorithm***

Paul Viola and Michael Jones introduced the Haar Cascade algorithm, a widely used method for real-time face detection. This algorithm is efficient and fast, making it suitable as a preprocessing step in emotion recognition systems. It detects faces from live video streams with minimal computational cost. However, its performance may degrade in complex backgrounds or low-light conditions, leading to false detections.

[5] ***Emotion Detection using FER-2013 Dataset***

Ian Goodfellow contributed to the development and use of the FER-2013 dataset, which contains thousands of labeled facial images representing various emotions. This dataset plays a crucial role in training deep learning models, particularly CNNs, for emotion classification tasks. While it improves model accuracy, challenges still exist in handling noise and real-world variations in input data.

[6] ***Emotion Recognition using DeepFace***

Facebook AI Research introduced DeepFace, a deep neural network-based system that combines facial recognition and emotion analysis. This approach achieves high accuracy by leveraging large-scale data and advanced neural network architectures. It enhances real-time performance and provides robust emotion detection capabilities, although it requires significant computational power and large datasets.

[7] ***Emotion Detection using Transfer Learning***

Various researchers have explored transfer learning techniques using pre-trained models such as VGG16, ResNet, and Inception. This approach reduces training time and improves model accuracy by leveraging knowledge from previously trained networks. Transfer learning is particularly useful for emotion recognition tasks with limited datasets, although fine-tuning is required for optimal performance.

[8] ***Multimodal Emotion Detection***

Research conducted by MIT focuses on multimodal emotion detection by combining facial expressions, speech, and textual data. This approach provides more accurate and reliable emotion recognition compared to single-modality systems. By integrating multiple data sources, the system can better understand human emotions. However, it increases system complexity and requires synchronization of different input modalities.

3 Proposed Methodology

The proposed system is an **Emotion Aware Human Computer Interaction (HCI) system using Real-Time Facial Expression Analysis** designed to overcome the limitations of traditional systems that lack emotional intelligence. The system focuses on enabling machines to understand human emotions by analyzing facial

expressions captured through a webcam in real time. It utilizes computer vision techniques for face detection and feature extraction, identifying key facial regions such as eyes, eyebrows, and mouth for analysis.

The system integrates a **Convolutional Neural Network (CNN)** model to automatically learn and classify emotions such as happy, sad, angry, surprised, and neutral with high accuracy. Unlike traditional approaches that rely on manual feature extraction, the proposed model uses deep learning to improve performance, adaptability, and robustness under different real-world conditions. The CNN processes extracted facial features and continuously enhances accuracy through training on large datasets.

Additionally, the system provides real-time emotion detection and instant feedback, making it highly responsive and user-friendly. This real-time capability reduces manual effort and enables natural interaction between humans and computers. The proposed system can be integrated into various applications such as online education, healthcare monitoring, virtual assistants, and interactive gaming, where understanding user emotions is essential.By leveraging deep learning and real-time processing, the system bridges the gap between human emotions and machine responses. It ensures more intelligent, adaptive, and personalized interactions, thereby improving user experience and making human-computer interaction more effective and natural.

3.1 System Architecture

The above architecture illustrates the overall design of the **Emotion Aware Human Computer Interaction System using Real-Time Facial Expression Analysis**. The system is structured into multiple layers including the User Layer, Frontend Layer, Processing Layer (Computer Vision + CNN), and Dataset Layer. It integrates real-time video processing with deep learning techniques to accurately detect and classify human emotions.

1. ***User Layer***

   The system primarily involves a single type of user:

   • ***User:*** *Interacts with the system through a webcam interface. The user's facial expressions are captured in real time, and the system provides emotion-based feedback dynamically.*

2. ***Frontend Layer (Streamlit / GUI Interface)***

   The frontend is developed using ***Python (Streamlit)****, providing a simple and interactive user interface. It includes the following components:*

   • **Webcam Interface:** Captures live video input from the user

   • **Display Screen:** Shows detected face with bounding box

   • **Emotion Output Panel:** Displays predicted emotion labels (Happy,Sad,Angry,etc.)

   • **Confidence Score Display:** Shows probability of detected emotion

   • **User Interaction Panel:** Allows system start/stop and monitoring

3. ***Processing Layer (Core System)***

   This is the core layer where all computation and intelligence are implemented. It consists of multiple modules:

• ***Face Detection Module***

This module detects human faces from live video frames using techniques such as **Haar Cascade Classifier**. It identifies and isolates facial regions by drawing bounding boxes, ensuring only relevant data is processed.

• **Image Preprocessing Module**

Captured facial images are preprocessed to improve quality and consistency. This includes:

- Conversion to grayscale
- Image resizing (e.g., 48×48 pixels)
- Normalization of pixel values
- This step reduces computational complexity and improves CNN performance.

• ***Feature Extraction Module***

In this stage, important facial features such as eyes, eyebrows, nose, and mouth are extracted. The CNN automatically learns these features through convolution layers, eliminating the need for manual feature extraction.

• ***CNN-Based Emotion Classification Module***

This is the core deep learning module where emotion classification occurs. The ***Convolutional Neural Network***

(CNN) processes extracted features and classifies emotions into categories such as:

- Happy
- Sad
- Angry
- Surprised
- Neutral
- The model assigns probabilities to each class and selects the emotion with the highest probability.

• ***Real-Time Output Module***

The final output is displayed in real time. The system continuously updates predictions and shows:

- Detected face with bounding box
- Predicted emotion label
- Confidence score

This ensures an interactive and user-friendly experience.

4. Dataset Layer

The system uses standard facial expression datasets such as:

•FER-2013Dataset

• CK+ Dataset

These datasets are used to train the CNN model. The images are preprocessed and labeled into different emotion categories to improve model accuracy and robustness

5. System Workflow

1. The user interacts with the system via webcam.
2. Live video frames are captured through the frontend interface.
3. Face detection module identifies facial regions.
4. Preprocessing is applied to enhance image quality.
5. Features are extracted using CNN layers.
6. The CNN model classifies the emotion.
7. The result (emotion + confidence score) is displayed in real time.

6. Key Features of the Architecture

• Real-Time Emotion Detection using webcam input

• Deep Learning-Based Classification (CNN) for high,accuracy

• Automated Feature Extraction (no manual intervention.required)

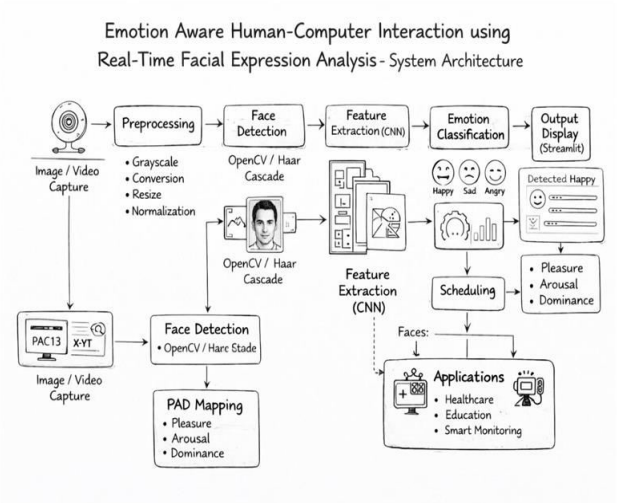
• Scalable and Modular Design for easy enhancement

• Efficient Image Processing for fast performance

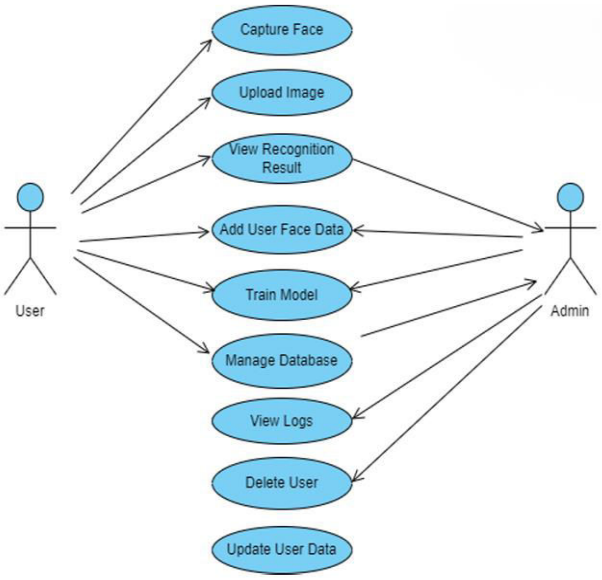
• Interactive User Interface using Streamlit

• Application Flexibility in healthcare, education, and smart systems

System Architecture diagram



3.2 Use Case diagram



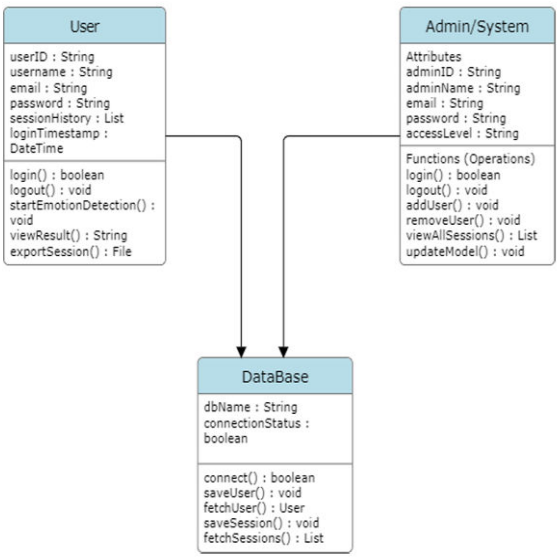
Use Case Diagram

The Use Case Diagram represents the interaction between different users and the Adaptive Learning

Platform with Gamified Quizzes & Analytics. It illustrates how various actors such as Student, Instructor, and Admin interact with the system functionalities. The Student can perform actions such as enrolling in courses, watching video lectures, taking adaptive quizzes, viewing progress, and earning rewards through gamification. The Instructor is responsible for managing courses, creating quizzes, and monitoring student performance. The Admin controls the overall system, including user management, system monitoring, and analytics.

3.3 Class diagram

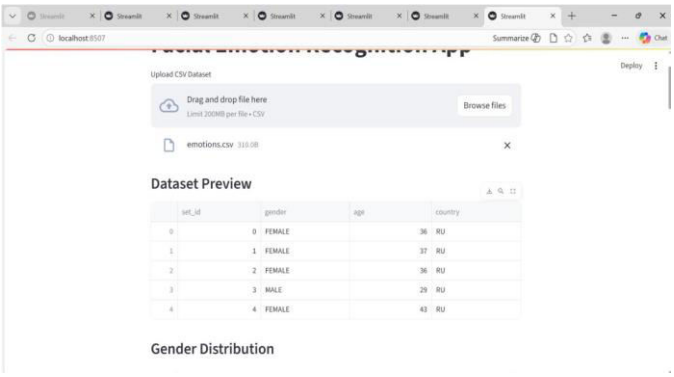
The class diagram represents the static structure of the system for "Emotions Aware Human-Computer Interaction Using Real-Time Facial Expressions Analysis". It shows the main classes, their attributes, methods, and relationships. Key classes include Camera,FaceDetector, EmotionRecognizer, UserInterface, and DataLogger. The diagram illustrates how the Camera captures video frames, which are processed by the FaceDetector to locate faces, analyzed by the EmotionRecognizer to determine emotions, and finally displayed through the UserInterface and recorded by the DataLogger. This diagram provides a clear overview of the system architecture and helps in understanding how different modules interact with each other.



Class Diagram

STATIC OUTPUT:

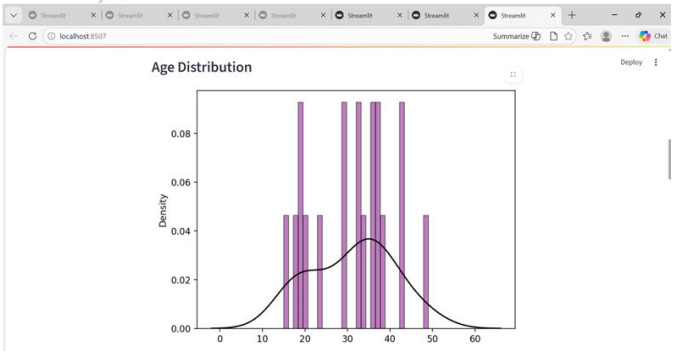
Dataset Preview:



Data Visualization:

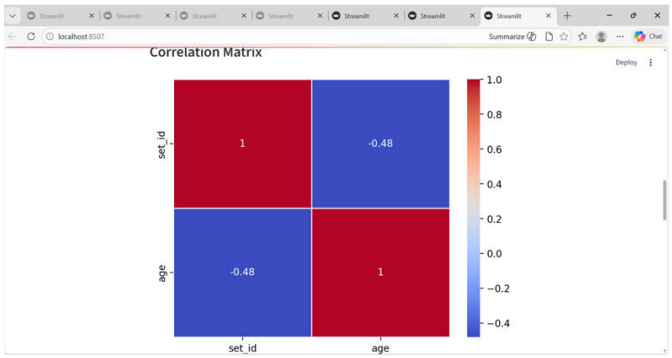
Gender Distribution

Age Distribution:

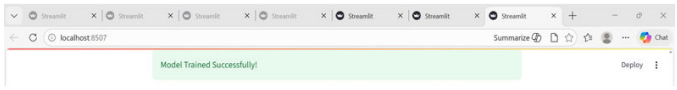


Correlation Matrix:

4.Results

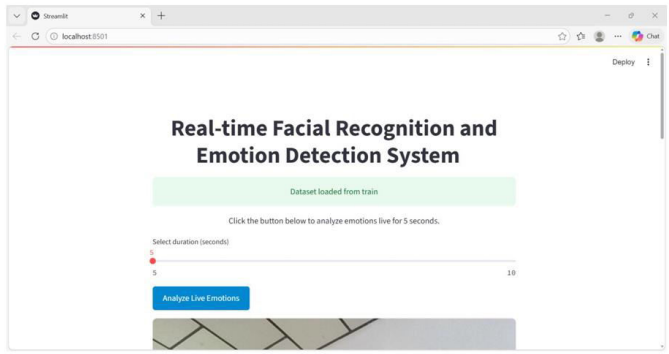


Model Training:

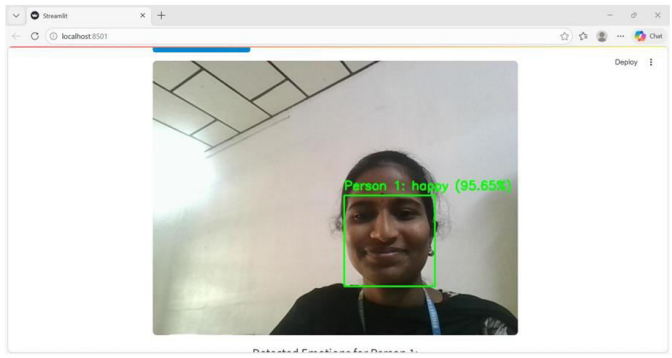


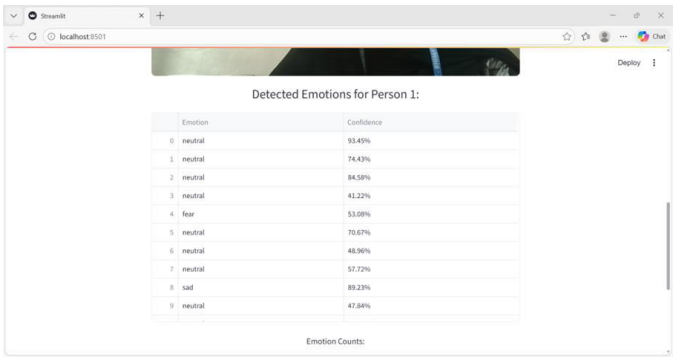
Sample Predictions

DYNAMIC OUTPUT:

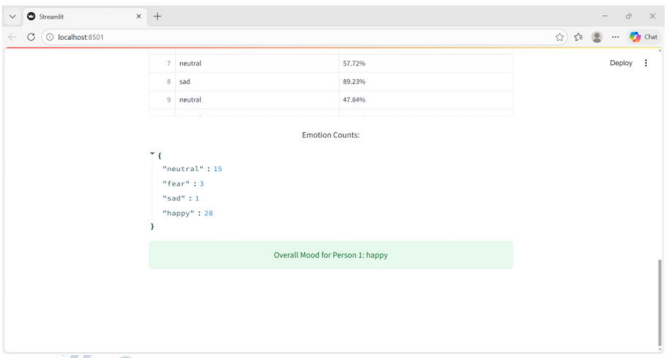


HAPPY

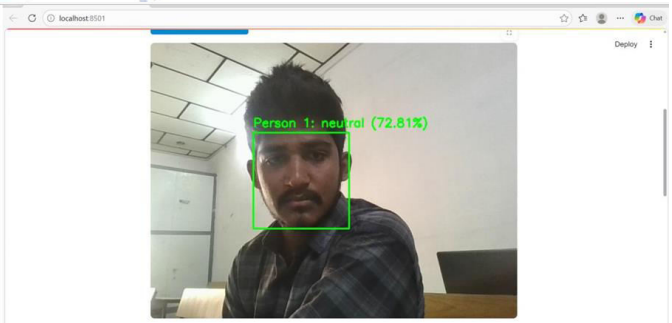


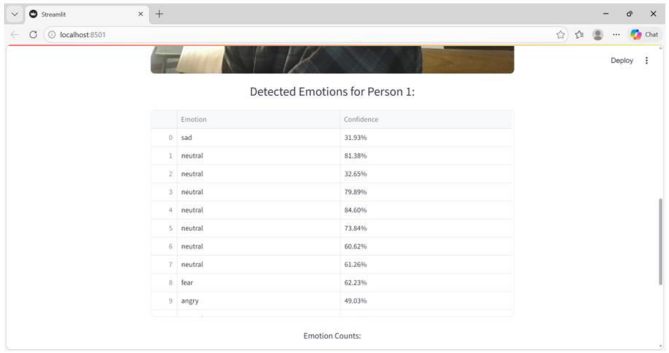


OVERALLMOOD

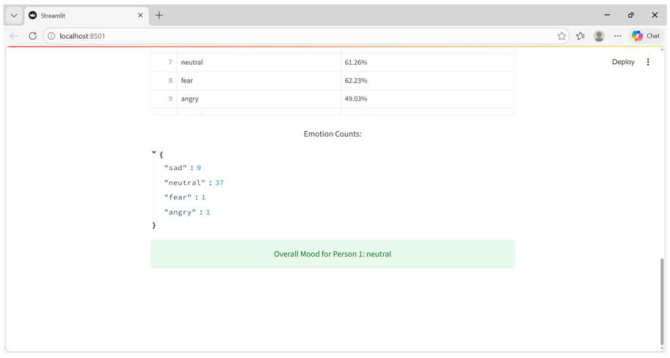


NEUTRAL:

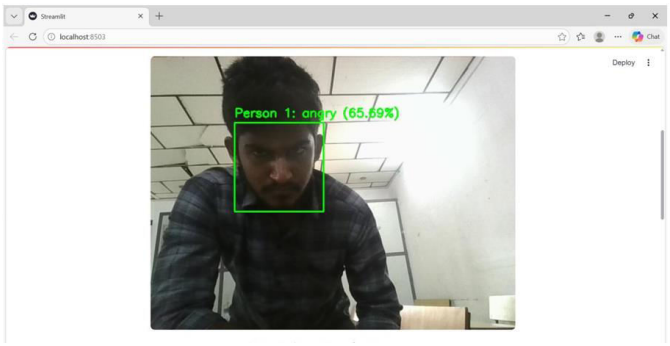


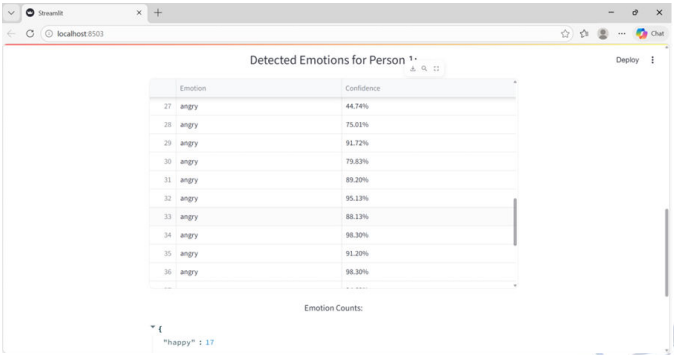


OVERALL MOOD:

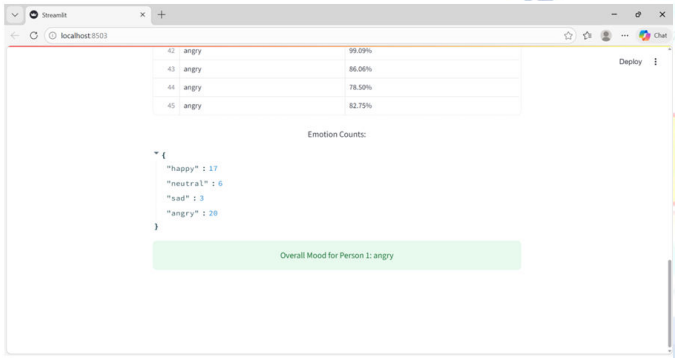


ANGRY:





OVERALL MOOD:



Conclusion

The Emotion Aware Human Computer Interaction system using real-time facial expression analysis is an effective and intelligent approach to improve the interaction between humans and computers. This project successfully demonstrates how computer vision and deep learning techniques can be used to detect and classify human emotions such as happy, sad, angry, and neutral from live facial inputs. By using tools like Python, OpenCV, and TensorFlow/Keras, the system is able to process facial features and provide accurate emotion recognition in real time. The implementation of this system highlights the importance of AI-based technologies in creating smart and responsive applications. It can be widely used in various fields such as healthcare for mental health monitoring, education for student engagement analysis, and in smart assistants to provide better user experience. Although the system performs well, its accuracy may vary depending on lighting conditions, facial clarity, and dataset limitations. Overall, this project proves that emotion-aware systems have great potential in future human-computer interaction, making technology more interactive, adaptive, and user-friendly.

Future Scope

In the future, the system can be enhanced by improving the accuracy of emotion detection using advanced deep learning models and larger datasets. It can also be extended to detect more emotions and support multiple languages. Integration with applications such as online learning, healthcare, and smart assistants can make the system more useful. Additionally, the system can be optimized to work on mobile devices and real-world environments with better performance.

- ***Improved Accuracy***

The system can be enhanced by using advanced deep learning models and larger datasets to achieve better emotion detection accuracy.

- ***More Emotion Detection***

The system can be extended to recognize a wider range of human emotions beyond the basic categories.

- ***Application Integration***

It can be integrated with applications such as online learning, healthcare, and smart assistants for better user interaction.

- ***Mobile Implementation***

The system can be optimized to run efficiently on mobile devices and embedded systems.

- ***Real-Time Performance***

Further improvements can be made to enhance real-time performance in different environmental conditions.

- ***Multi-Face Analysis***

The system can be improved to handle multiple faces more efficiently and provide accurate emotion detection for each individual.

- ***Voice-Based Emotion Detection***

Future enhancement can include speech analysis to detect emotions from voice along with facial expressions.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] P. Ekman and W. V. Friesen, Facial Action Coding System: A Technique for the Measurement of Facial Movement, Consulting Psychologists Press, 1978.

[2] I.Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.

[3] OpenCV Documentation, Open Source Computer Vision Library. Available: https://opencv.org/

[4] TensorFlow Documentation, An End-to-End Open Source Machine Learning Platform. Available: https://www.tensorflow.org/

[5] Zeng, Z., Pantic, M., Roisman, G. I., & Huang, T. S. (2009). A survey of affect recognition methods: Audio, visual, and spontaneous expressions. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(1), 39–58.

[6] Li, S., & Deng, W. (2020). Deep Facial Expression Recognition: A Survey. IEEE Transactions on Affective Computing, 13(1), 119–135.

[7] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[8] OpenCV Documentation – https://docs.opencv.org/

[9] TensorFlow & Keras Documentation – https://www.tensorflow.org/

[10] Datasets: FER2013, CK+, AffectNet – Publicly available facial expression datasets used for model training and evaluation.

[11] Darwin, C. The Expression of Emotions in Man and Animals; Murray: London, UK, 1873.

[12] Ekman,P. Emotion in the Human Face; Cambridge University Press: Cambridge, UK, 1982.

[13] Plutchik, R. Emotions: A Psychoevolutionary Synthesis; Harper & Row: New York, NY, USA, 1980.

[14] Whissell, C. The dictionary of affect in language. In Emotion: Theory, Research, and Experience, Volume 4: The Measurement of Emotions; Plutchik, R., Kellerman, H., Eds.; Academic Press: New York, NY, USA, 1989; Chapter 5, pp. 113–131.

[15] Russell, J.A. Measures of emotion. In Emotion: Theory, Research, and Experience. The Measurement of Emotions; Plutchik, R., Kellerman, H., Eds.; Academic Press: New York, NY, USA, 1989; Volume 4, Chapter 4, pp. 83–111. 24. Cochrane, T. 8 dimensions for the emotions. Soc. Sci. Inf. Spec. Issue The Lang. Emot. Concept. Cult. Issues 2009, 48, 379–420. 25. Sedikides, C. Changes in the Valence of the Self as a Function of Mood. Rev. Personal. Soc. Psychol. 1992, 14, 271–31

[16] Neumann, R.; Seibt, B.; Strack, F. The influence of mood on the intensity of emotional responses: Disentangling feeling and knowing. Cogn. Emot. 2001, 15, 725–747. [CrossRef]

[17] Mehrabian, A. Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. Curr. Psychol. 1996, 14, 261–292. [CrossRef]

[18] Arifin, S.; Cheung, P.Y.K. A computation method for video segmentation utilizing the pleasure-arousal-dominance emotional information. In Proceedings of the MM '07: Proceedings of the 15th ACM International Conference on Multimedia , Augsburg, Bavaria, Germany, 23–28 September 2007; pp. 68–77. [CrossRef]

[19] Arellano, D.; Perales, F.J.; Varona, J. Mood and Its Mapping onto Facial Expressions. In Proceedings of the Articulated Motion and Deformable Objects; Perales, F.J., Santos-Victor, J., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 31–40.Biehn, C. Facial Expression Repertoire (FER). 2005. Available online: http://research.animationsinstitut.de/ (accessed on 18 January 2014)

[20] Kanade, T.; Cohn, J.F.; Tian, Y. Comprehensive database for facial expression analysis. In Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580), Grenoble, France, 28–30 March 2000; pp. 46–53.

[21] Lucey, P.; Cohn, J.F.; Kanade, T.; Saragih, J.; Ambadar, Z.; Matthews, I. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, San Francisco, CA, USA, 13–18 June 2010; pp. 94–101.

[22] Saragih, J.M.; Lucey, S.; Cohn, J.F. Deformable model fitting by regularized landmark mean-shift. Int. J. Comput. Vis. 2011, 91, 200–215. [CrossRef]

[23] Valstar, M.; Pantic, M. Induced Disgust, Happiness and Surprise: An Addition to the Mmi Facial Expression Database. In Proc. 3rd Intern. Workshop on EMOTION (Satellite of LREC): Corpora for Research on Emotion and Affect; 2010; p. 65. Available online: https://ibug.doc.ic.ac.uk/research/mmi-database/ (accessed on 4 July 2024).