Innovation and Integrative Research Center Journal
ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-5 | May-2026 | Page 798-808

EMO VISION: REAL- TIME FACIAL EMOTION RECOGNITION SYSTEM

[1]Prakhar Shukla, [2]Dr. Anjali Chandra, [3]Deepanshu Sahu
[1,3]UG- Computer Science & Engineering with Specialization in Artificial Intelligence, [2]Assistant Professor
Shri Shankaracharya Institute of Professional Management and Technology, Raipur, Chhattisgarh, India

Abstract

Human emotions play a vital role in communication, influencing decision-making, learning ability, behaviour, mental health, and interpersonal interactions. Emotion recognition has become increasingly important in fields such as online education, healthcare, security, customer experience, robotics, and human–computer interaction. Traditional methods like surveys, manual observation, and physiological tests are subjective, time-consuming, and unsuitable for continuous real-time analysis. Advances in Artificial Intelligence (AI), Machine Learning (ML), and Computer Vision (CV) have enabled the development of automated and non-intrusive emotion detection systems.

This project, Intelligent Real-Time Emotion Recognition Using Computer Vision and Deep Learning, aims to develop a real-time system that detects facial expressions through a webcam, classifies emotions using a pre-trained deep learning model, and logs results for behavioural analysis. The system uses OpenCV Haar cascade for face detection and Deep Face for emotion classification. Emotions such as happiness, sadness, anger, fear, disgust, surprise, and neutrality are identified with high accuracy. Results are stored with timestamps and analysed to generate insights, making the system useful for applications in education, mental health, and behavioural research.

Keywords: Facial Emotion Recognition, Deep Learning, Convolutional Neural Networks (CNN), Computer Vision, Artificial Intelligence (AI), Machine Learning (ML), OpenCV, Deep Face, Real-Time Emotion Detection, Behavioural Analysis, Emotion Logging, Human–Computer Interaction, Learning Analytics.

1. INTRODUCTION

Emotions are non-verbal communication, which is a potent tool of human behavior, thought processes and interrelation [1]. This is because the capability to process such emotional messages is valuable not only to humans but also the intelligent systems that wish to engage with people in a natural manner. Since digital platforms are increasingly adopted in education, entertainment, healthcare and business, machine interpretation of emotional feedback becomes ever more pertinent. The AIs have allowed computers to identify faces, read expressions and classify emotions in an extraordinarily accurate manner to allow new applications in affective computing [2][3]. Historically, emotion recognition had to be implemented by hand by an expert or through self-reports or physiological measurements like EEG, GSR, or ECG monitors. Despite its effectiveness, these methodologies are limited by the presence of human bias, inconvenience, intrusiveness and failure to continuously monitor the emotions [4]. Computer Vision methods greatly alleviate these drawbacks since they enable the system to detect emotions all they do is

ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-5 | May-2026 | Page 798-808

to scan a facial expression using a camera. Facial expression being a universal expression of emotional conditions, it is a good source of automated classification systems [1][5]. The Intelligent Real-Time Emotion Recognition Using Computer Vision and Deep Learning system is designed to incorporate the face-detection techniques and deep learning-driven emotion-classification models to produce a smooth real-time solution. The system takes live video with the aid of a webcam, identifies facial areas with the help of pre-trained Haar cascade classifiers [6], and processes these areas with the help of a Deep Face model identifying a predominant emotional state [7]. It is an everlasting system that records all the emotional predictions in a CSV file with time stamps. The analysis module can be later run by the user to produce insights of any significance regarding to emotional trends during a session. This project is in line with contemporary requirements in many areas. Teachers in education and intelligent tutoring system can track the level of student engagement. In the medical field, therapists are able to examine emotional variation in the patient's receiving treatment. Organizations are able to assess customer response in customer service. The abnormal emotional patterns can assist in detecting the possible threats in security settings. The suggested system is easy, flexible, and open, which is why it can be applied to different real-life situations.

2. LITERATURE REVIEW

2.1. History of Facial Emotion Recognition Systems

Facial Emotion Recognition (FER) is a recent field in the study of Artificial Intelligence, especially in Computer Vision and Deep Learning. The new developments have made it possible to create automated systems that could capture human emotions based on facial expressions with better accuracy and efficiency. Kopalidis et al. [1] have provided a thorough overview of deep learning-based FER approaches, including popular datasets, including FER2013, RAF-DB, and Affect Net, metrics of evaluation, and model architecture. Their paper identifies the main issues such as occlusion, lighting differences, and cross-dataset generalization and the importance of the robust and flexible FER systems.

2.2. FER Models Limitations of Deep Learning

A number of investigations have been done on enhancing the performance of a model through advanced methods. Raju et al. [2] suggested an emotion prediction system based on deep learning in a distributed federated learning with SMOTE oversampling to solve the imbalance of classes. Although this method improves the accuracy of the model, it adds more computational complexity and reliance on a distributed infrastructure. In the same manner, Bhagat et al. [3] adopted a Convolutional Neural Network (CNN) in recognizing emotions of video inputs. The model was also limited when working with real world conditions, including different illumination and the different facial orientations although it has attained satisfactory classification results.

2.3. Dataset use in Emotion Recognition

The quality of datasets is very vital in enhancing the performance and generalization of FER systems. Pourramezan Fard et al. [4] proposed Affect Net+, a larger dataset with soft-label annotations to make the models more reliable and effective to train. Such a contribution plays a major role in enhancing the

ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-5 | May-2026 | Page 798-808

quality of data sets and generalization of the models but does not help in the problems associated with the real-time system implementation or the actual implementation.

2.4. Multimodal and Comparative Approaches

Multimodal studies have been conducted recently to improve the accuracy of emotion recognition. Jena et al. [5] created a deep learning model of negative speech emotion recognition, stating the significance of using a combination of various modalities like facial expression and speech signals. Moreover, Joshi et al. [6] analysed the comparative study of different deep learning architectures used to learn FER using different data sets. Their discussion contains many details of the strengths and weaknesses of various models but does not focus on the real-time system integration and implementation.

2.5. Real-Time Implementation Problems

Pradeep et al. [7] suggested a real-time facial emotion detection system which involves CNN and OpenCV and explained that it can be practically implemented with the help of webcam input. Although their system is effective in emotion classification, it is more of a detection system and does not have capabilities of emotion logging and analytical visualisation to analyse behaviour. This means that there exists a loophole in the current systems, as real-time detection is realized but extensive data analysis is not always present.

2.6. Research Gap And Proposed Approach

Based on the analysed literature, it becomes clear that the majority of the existing research is devoted to enhancing the accuracy of classification or the quality of the data or the performance of the models. Nonetheless, little focus has been directed towards the creation of lightweight and real-time systems that combine emotion recognition with logging and analysis functionalities. The proposed system fills these gaps by offering a feasible and affordable solution, which integrates the real-time emotion identification with the behavioural data analysis, which would be applicable to real-world application.

3. METHODOLOGY

3.1. System Architecture and Design

The proposed system, Intelligent Real-Time Emotion Recognition Using Computer Vision and Deep Learning, is developed with a modular and layered architecture, which includes real-time video processing, face detection, deep learning-based emotion recognition, data logging, and analysis visualization. The architecture is designed to be efficient, scalable and run-on general-purpose hardware without the need of specialized computational hardware.

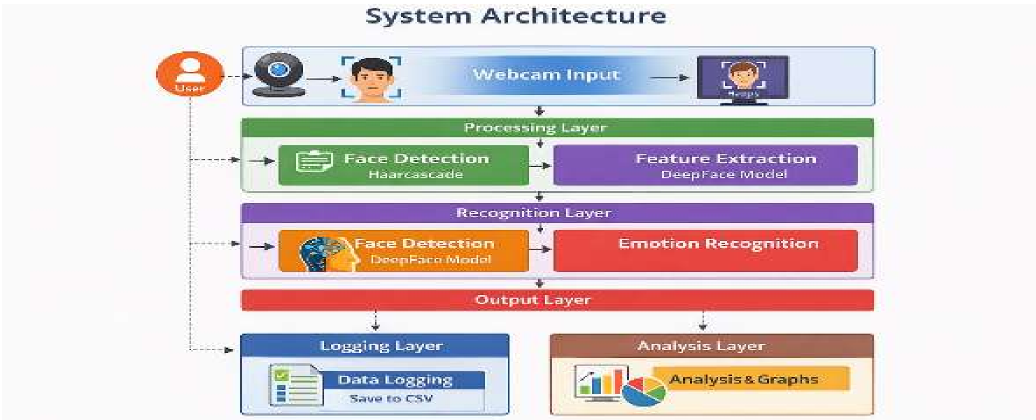

Figure 1: System Architecture

ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-5 | May-2026 | Page 798-808

The system is divided into several functional layers: the Input Layer, Processing Layer, Recognition Layer, Logging Layer and Analysis Layer. The processing of the video starts with the acquisition of the video with help of a web camera by using the OpenCV library. Every frame that is captured is considered as a separate input and handled sequentially to ensure that the system continues to run in real-time. As shown in Figure 1. Processing Layer face is detected with the help of Haar cascade classifier through Viola-Jones algorithm. The Recognition Layer operates the Deep Face framework that is based on Convolutional Neural Networks (CNNs) to identify emotions. The Logging Layer captures the identified emotions and timestamps in an organised CSV file, and the Analysis Layer then uses this data to create statistical information and visualisations.
This multilayered structure guarantees high maintainability, efficient data flow, low latency and reliable real-time performance.

3.2. System Design

The system design specifies the workflow, interaction between modules, and the flow of data within the application. The algorithm starts with the real-time frame capture via webcam interface. The frame is converted to grayscale and faces are identified to extract the region of interest (ROI).
The extracted face is then pre-processed to fit the input needs of the deep learning model. Resizing and normalization are part of preprocessing to maintain uniformity when representing data. The processed image is sent to Deep Face model where it is classified as an emotion.
The system detects emotions like:

- Happiness
- Sadness
- Anger
- Fear
- Surprise
- Disgust
- Neutral

Each prediction has the dominating emotion, which is shown in real-time on the video feed with the help of bounding boxes and labels.

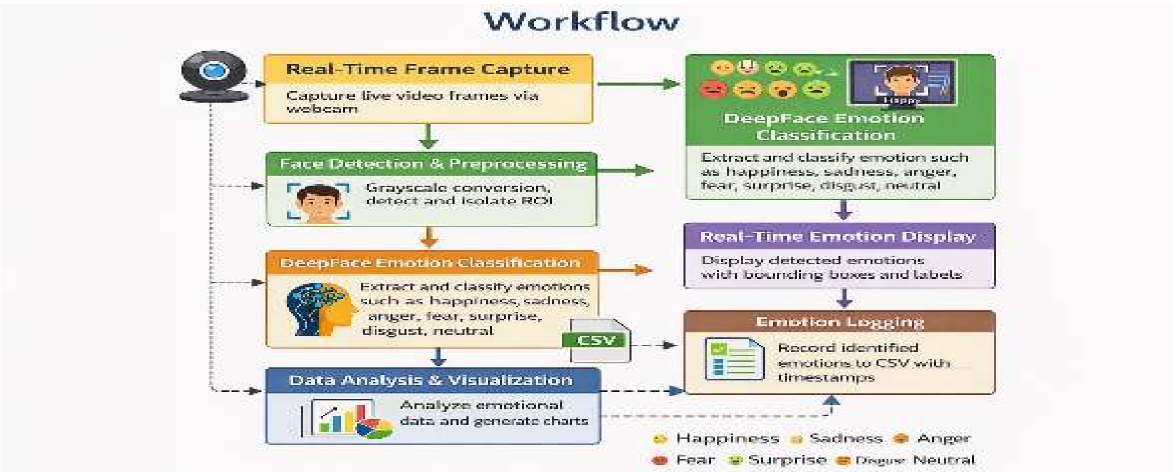

Figure 2: Flowchart – system Design

ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-5 | May-2026 | Page 798-808

It is a continuous loop system with the processing of the frames going in a sequence. Identified emotions are recorded to a CSV file and timestamps. This makes it possible to chronologically follow emotional changes during the session. As shown in Figure 2.

The analysis module reads the data that has been stored and calculates important measures like the frequency and percentage distribution of emotions, and duration of the session. Emotional patterns are visualized by creating graphical representations like bar charts.

3.3. Modules Description

• Input Module

Input Module takes live video with a webcam with the help of OpenCV. It guarantees the continuity of frame capture and consistent real time operation.

• Face Detection Module

This module identifies facial regions with Haar cascade classifier. It isolates areas of interest (ROI) and disregards the background noise enhancing processing efficiency.

• Preprocessing Module

The preprocessing module preprocesses the extracted facial region to be used as model input. It includes resizing and normalization, ensuring consistency and improved classification accuracy.

• Emotion Recognition Module

The Deep Face framework applied to classify emotions is based on CNNs in this module. It recognizes the prevailing emotional expression in each face detected.

• Display Module

The display module gives the system interactivity and usability by overlaying bounding boxes and emotion labels on the video feed, which also gives real-time visual feedback.

• Logging Module

Logging module logs identified emotions with timestamps in a CSV file. It provides orderly and sequential storage of data to be further analysed.

• Analysis Module

The analytics module is used to work with logged data in Pandas and Matplotlib. It creates statistical information like distribution of frequencies and graphs.

3.4. Algorithms and Technologies Used

• Deep Learning (Emotion Recognition with CNN)

The system makes use of Deep Face, Convolutional Neural Networks (CNNs) based system to extract facial features and classify emotions. The use of pre-trained models ensures high accuracy and eliminates the need for extensive training.

• Face Detection Algorithm

Haar cascade classifier is an algorithm of the Viola Jones that is employed in efficient and fast face detection in real-time applications.

• Data Processing Algorithms

Deterministic logic (Python) is used to perform statistical calculations like frequency analysis, percentage calculation, and session duration.

• Programming and Frameworks

ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-5 | May-2026 | Page 798-808

The system is based on Python and libraries OpenCV, Deep Face, Pandas, and Matplotlib. These tools offer effective processing of images, model assimilation and data analysis.

• Data Storage

A CSV-based storage mechanism is used to log emotions and timestamps. This guarantees ease, reliability and simplicity of accessing data.

3.5. System Workflow

The overall workflow of the system is defined as follows:

Webcam Input → Face Detection → Preprocessing → Emotion Recognition → Display Output Data Logging Data Analysis Visualization.

*3.6.*Performance Things and Benefits

The system should be efficient in real-time and with low latency and optimal resource usage. It is stable to moderate changes in lighting and face position. Lightweight models and open-source tools will be used to make the system cost-effective and affordable.

The suggested methodology is a compromise between accuracy and efficiency and is easy to deploy. It is conducive to real-world uses in education, health care, behavioural analysis and human-computer interaction. The system is also flexible to future upgrades i.e. multi-face recognition, superior visualization displays and cloud computing.

4. RESULT

4.1. Introduction

The chapter includes the assessment of Intelligent Real-Time Emotion Recognition System with the help of Computer Vision and Deep Learning. A real-time test with the usage of standard webcam was applied to the system to test how it can detect the faces, identify emotions correctly and have a structured log to do further analysis. The primary goal is to measure the performance of the system regarding real-time responsiveness, the reliability of categories, logging, and visualization. The findings indicate the usefulness of combining the computer vision and deep learning technologies to real-life applications of emotion recognition.

4.2. Experimental Setup

It was written in Python and was tested on a computer with no graphics hardware. The input device was a conventional webcam, which was used to record live video streams. Video capture and face detection were done using OpenCV and emotion classification was done using Deep Face framework. Images were constantly taken by the system and analysed in real time. The analysis was performed on the individual frames of the video to identify the facial expression and the classification of the emotion with a minimum delay. Testing was done in different light conditions and movement of the users to test on robustness. The arrangement was such that the system could be used effectively in real life settings like in classrooms, offices and personal areas

Innovation and Integrative Research Center Journal
ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-5 | May-2026 | Page 798-808

4.3. Results of Emotion Detection In Real-Time

The system was able to recognize faces and identify emotions in real time. The Haar cascade classifier was used to locate facial regions which was then inputted to the Deep Face model to analyze the emotion. A predicted emotion was shown on the screen and a bounding box was enclosed around the identified face, which gave immediate visual feedback. In the process of testing, the system was correct in the recognition of emotions like happiness, sadness, anger, fear, surprise, disgust and neutral. The detection process was able to stabilize with the normal lighting condition with little delay between the capture of the frame and output display. The system was very adaptable to the variation on the facial expressions and the predictions are dynamic which is evidence of the real-time performance of the system and responsiveness. As shown in Figure 3. WaIT implementation has enhanced the effectiveness and efficiency of the customer service process within an organization, as demonstrated below:

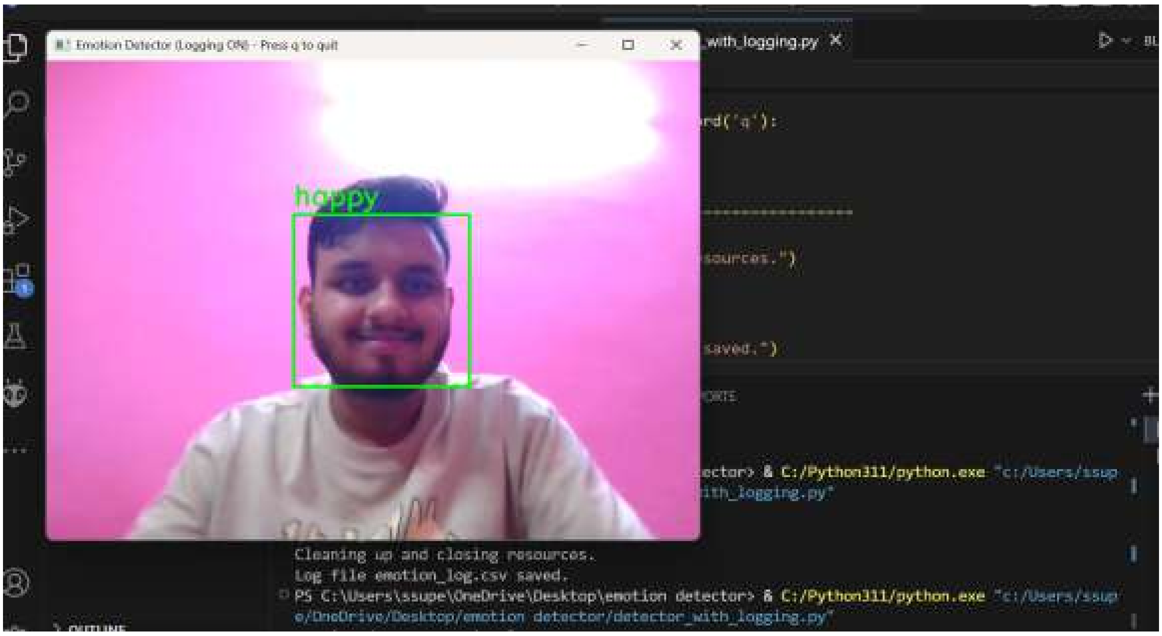

Figure 3: Real-Time Emotion Detection Result

4.4. Emotion Logging Results

The system has an effective logging system that tracks every emotion detected and a timeline in a CSV file. The process is successful in the sense that, all the emotional changes will be saved in chronological order throughout the session. The resulting log file is a structured level of data of emotional change with time. It can be applied to behavioral analysis, learning analytics and mental health monitoring. It was also discovered that the logging system was reliable and the timestamps were accurate and no data were lost during the execution. Minor emotional shifts were also well-captured, which could be analyzed in detail in the course of the sessions.

4.5. Statistical Analysis of Emotions

The analytics module assimilates the data log and derives meaningful information. It calculates the overall detections, the frequency of each emotion, distribution in the percentage and duration of the session. The findings prove that emotional patterns are captured well by the system. Through frequency and percentage distribution, one can identify the common emotions throughout a session. Calculation of the session duration also assists in interpretation of trends of emotions with time. As shown in Figure 4. This analysis will give an overview on the user behavior and emotional states.

```
Loading data from 'emotion_log.csv'...
Data loaded successfully. Analyzing...

--- Session Analysis ---
Total Detections: 421
Session Duration: 22.81 seconds
Detections per Second: 18.46

--- Emotion Frequency ---
- neutral: 261 times (62.0%)
- happy: 85 times (20.2%)
- sad: 30 times (7.1%)
- fear: 27 times (6.4%)
- angry: 17 times (4.0%)
- surprise: 1 times (0.2%)

Generating analysis plot...
Analysis plot saved as 'emotion_analysis.png'
```

Figure. 4: Statical Analysis

*4.6.*Graphical Representation

The system uses Matplotlib to produce graphical visualizations so that they can be more easily interpreted. The frequency of the identified emotions is plotted into a bar chart. The height of each of them is associated with the frequency of occurrence of a certain emotion, and the bar can be identified with it. Percentage values also exist which are displayed in a better way. The graphic representation makes it easy to study the tendencies of emotions and uses it to discover the prevailing emotions and behaviour patterns. As shown in Figure 5. This is a capability that improves usability of the system since raw data is converted into easily readable format.

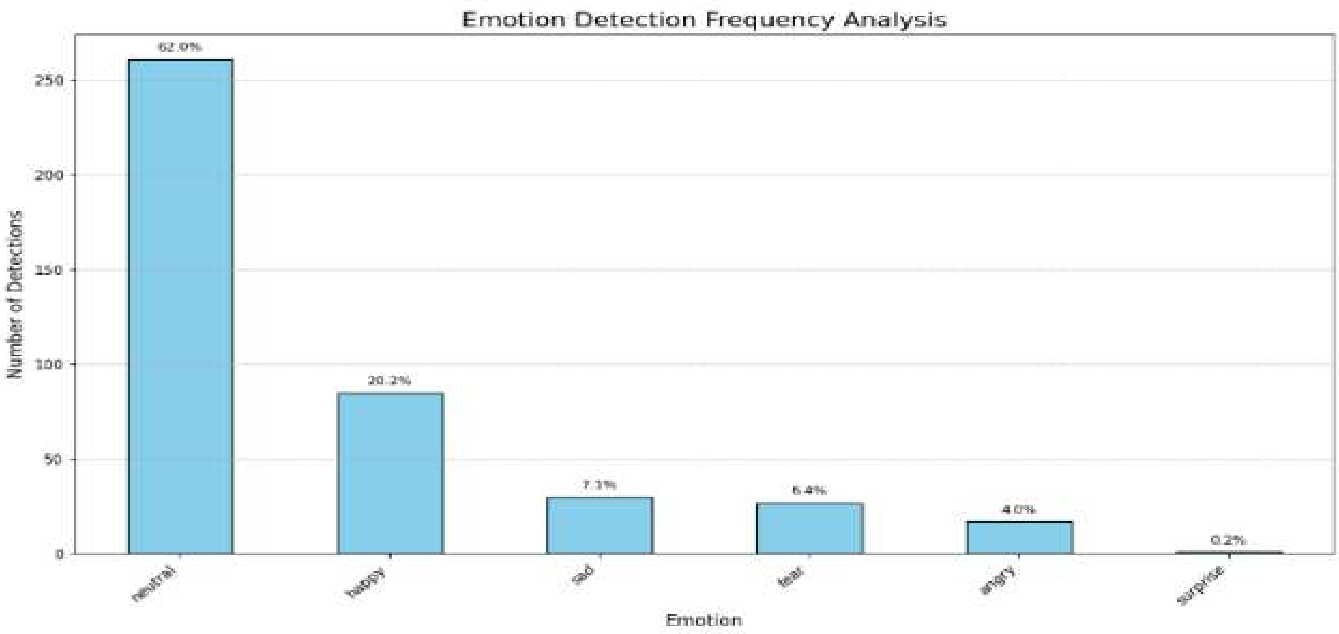

Figure 5: Figure of Frequency of emotion

ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-5 | May-2026 | Page 798-808

4.7. Performance Evaluation

The system was tested in terms of real time processing, accuracy, computational efficiency and strength. It exhibited live performance with little latency between input capture and output display. Haar cascade classifier was used to detect faces very quickly, whereas Deep Face model provided credible results in terms of emotion identification. The system was lightweight and could be run easily on CPU, without having to use special hardware. It worked satisfactorily in a medium light situation and minimal facial movements. However, the performance dropped slightly when the conditions became challenging like in dark lighting, occlusion, or a moving object.

5. DISCUSSION

5.1. Strengths of the System that is Implemented

The overall assessment of the Real-Time Emotion Detection and Analytical Logging System shows that the system has a number of important strengths in the architecture and performance. Incorporating Deep Face inference and OpenCV localization, the system will provide a healthy compromise between detection rate and real-time processing.

•Temporal Data Integrity:

In contrast to typical detection scripts, this system uses high-precision time stamps (within milliseconds), so even fast micro-expressions are recorded and stored to be analysed.

•Automated Statistical Reporting:

Due to the implementation of analysis.py, there is no longer any need to enter data manually because the duration of each session, the number of detections and the percentage of frequency are automatically computed.

•Resource Optimization:

With the help of grayscale conversion and the Haar Cascade classifiers to localize faces first, the system minimizes its computational load, which means that the system can be implemented in a regular hardware without the need of a dedicated GPU.

Visual Clarity: The automated creation of the emotion_analysis.png plot offers a real time, quantitative summary of the affective state of the user and thus the data can be deciphered by the non-technical stakeholders.

5.2. Limitations and Technical Problems

No matter how powerful the system is, it is not without some limitations typical of the computer vision and deep learning sphere.

Filter: The Haar Cascade and Deep Face models are quite environmental sensitive and the quality of the input video can significantly affect the accuracy of the results, and it can also be negatively affected by low-light conditions or extreme camera angles.

Contextual ambiguity: Although the system can recognize the dominant emotion using the facial geometry, it might not be able to distinguish between real expressions and certain posed or culturally specific facial expressions.

•Computational Latency: The computational cost of complex ensemble models to detect emotions with high accuracy is potentially high, which may limit the number of Detections per Second of less powerful devices.

•Interpretability Concerns: Deep learning models are usually a black box, and it is difficult to see how a particular emotion classification in a frame of uncertainty works.

5.3. Practical Applications

The applications of this real-time emotional monitoring system have implications in several high-impact areas.

Mental Health Monitoring: The system will be applicable to record emotional trends throughout the therapy sessions that will give therapists objective data of patient progress and emotional triggers.

•User Experience (UX) Research: The logging and analysis tools provide companies with the ability to evaluate customer responses to software interfaces or ads in real-time, to determine areas of frustration or joy.

•Technology in Education: In distance learning classes, teachers have the ability to check the level of engagement of students and modify the teaching methods according to the emotional frequency of the lesson.

Human-Computer Interaction (HCI): This study helps to create emotionally-intelligent AI, including virtual assistants capable of adjusting their tone and behaviour according to the mood of the user at that moment.

6. CONCLUSION

The Intelligent Real-Time Emotion Recognition Using Computer Vision and Deep Learning project is a successful example of how an effective and functional system to detect and analyse human emotions in real time can be developed. The system combines deep learning models with computer vision to recognize face expressions and categorize emotions with high precision with an ordinary web camera. The combination of OpenCV face detection and Deep Face emotion recognition framework allows the system to strike a balance between performance, accuracy, and computation efficiency.

The most important success of this project is the possibility of real-time detection of emotions without the need to use special equipment, like GPUs. The system is easy to run-on general-purpose devices, thus affordable and efficient. Moreover, a logging system can be used to record emotional states continuously and provide timestamps to track behaviour in a session. The system can also be improved with the introduction of an analysis module which can give statistical information and graphical view of emotional patterns.

Its ability to detect emotions like happiness, sadness, anger, fear, surprise, disgust, and neutrality with reliable performances in normal circumstances is proven by the results that the system produces. The system is stable, reactive and user-friendly, and is applicable in the real world. In general, the project has shed light on the ability of combining deep learning with computer vision to develop smart systems that can learn human feelings in an unobtrusive and automated way.

References

[1] Kopalidis, T., Solachidis, V., Vretos, N., & Daras, P. (2024). Advances in facial expression recognition: A survey of methods, benchmarks, models, and datasets. Information, 15(3), 135.
[2] Raju, V. N., Saravanakumar, R., Yusuf, N., Pradhan, R., Hamdi, H., Saravanan, K. A., ... & Askar, M. A. (2024). Enhancing emotion prediction using deep learning and distributed

ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-5 | May-2026 | Page 798-808

federated systems with SMOTE oversampling technique. Alexandria Engineering Journal, 108, 498-508.

[3] Bhagat, D., Vakil, A., Gupta, R. K., & Kumar, A. (2024). Facial emotion recognition (FER) using convolutional neural network (CNN). Procedia Computer Science, 235, 2079-2089.

[4] Pourramezan Fard, A., Hosseini, M. M., Sweeny, T. D., & Mahoor, M. H. (2024). AffectNet+: A Database for Enhancing Facial Expression Recognition with Soft-Labels. arXiv e-prints, arXiv-2410.

[5] Jena, S., Basak, S., Agrawal, H., Saini, B., Gite, S., Kotecha, K., & Alfarhood, S. (2025). Developing a negative speech emotion recognition model for safety systems using deep learning. Journal of Big Data, 12(1), 54.

[6] Joshi, R. C., Juyal, A., Mishra, A., Verma, A., & Singla, K. (2024). Deep learning-based face emotion recognition: A comparative study. International Journal of Performability Engineering, 20(1), 1.

[7] Pradeep, V., Sumukha, B. S., Richards, G. R., & Prashant, S. P. (2024, April). Facial emotion detection using cnn and opencv. In 2024 International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS) (pp. 1-6). IEEE.

[8] Serengil, S. I., & Ozpinar, A. (2020, October). Lightface: A hybrid deep face recognition framework. In 2020 innovations in intelligent systems and applications conference (ASYU) (pp. 1-5). IEEE.

[9] Viola, P., & Jones, M. (2001, December). Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001 (Vol. 1, pp. I-I). Ieee.

[10] Bradski, G. (2000). The opencv library. Dr. Dobb's Journal: Software Tools for the Professional Programmer, 25(11), 120-123.

[11] McKinney, W. (2010). Data structures for statistical computing in Python. scipy, 445(1), 51-56.

[12] Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. Computing in science & engineering, 9(03), 90-95.

[13] Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. Journal of personality and social psychology, 17(2), 124.

[14] Dhall, A., Goecke, R., Gedeon, T., & Sebe, N. (2016). Emotion recognition in the wild. Journal on Multimodal User Interfaces, 10(2), 95-97.