iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.80838

Mental Health Monitoring System

Miss. Shreya Bhise[1], Miss. Shreya Rane[2], Miss. Meghana Shivaprasad[3], Prof. Shraddha Katkar[4]
Department of Computer Science, MIT Art, Design and Technology, University, Pune

Abstract: Mental health challenges are increasing globally, while access to timely and affordable support remains limited. This paper presents an AI-based Mental Health Monitoring System that provides real-time emotional analysis and adaptive support using a multimodal approach. The system integrates facial emotion recognition using Convolutional Neural Networks (CNN), voice tone analysis using Web Audio APIs, and text sentiment analysis through Natural Language Processing techniques.

The system has evolved from an initial prototype into a full-stack platform incorporating secure user authentication, session lifecycle management, journaling, and emotion analytics. It follows a client-server architecture where backend services handle processing and storage using a PostgreSQL database. The platform is accessible via web and mobile browsers through Progressive Web App (PWA) support.

The system demonstrates real-time performance with low latency and provides context-aware responses through an AI-driven chatbot. This solution serves as an assistive tool for emotional awareness and early intervention.

Keywords: Mental Health Monitoring, Multimodal Emotion Detection, CNN, Sentiment Analysis, Conversational AI, PWA.

I. INTRODUCTION

Mental health disorders such as anxiety, depression, and stress significantly impact individuals' well-being and productivity. Despite increasing awareness, many individuals hesitate to seek professional help due to stigma, cost, and limited accessibility.

Artificial Intelligence (AI) enables the development of systems capable of detecting emotional states and providing timely support. Emotion recognition using facial expressions, voice tone, and textual input plays a crucial role in such systems.

This work presents an implemented AI-based mental health monitoring system that combines multimodal emotion detection with an interactive chatbot. The system provides real-time emotional insights, session tracking, and analytics, enabling proactive emotional well-being management.

II. LITERATURE REVIEW

Recent advancements in affective computing and mental health technologies have leveraged deep learning, multimodal fusion, and AI-driven platforms. This section summarizes key contributions from recent literature.

A. Multimodal and Transformer-Based EEG Emotion Recognition (2025–2024)

Zhang et al. [1] reviewed EEG-based multimodal fusion techniques, highlighting improved accuracy through sensor integration while noting challenges such as complexity and synchronization. Zhu et al. [2] proposed SOGPCN, a graph-based pseudo-3D CNN achieving high accuracy in emotion recognition. Pan et al. [3] introduced a dual-attention transformer for long-term EEG analysis, improving contextual understanding but increasing computational cost.

Chu et al. [4] developed multimodal conversational systems combining EEG, speech, and text. Lee et al. [5] introduced a synchronized EEG-audio-video dataset for realistic emotion recognition research.

B. Deep Learning Models and AI in Therapy (2023–2022)

Yuvaraj et al. [6] applied 3D-CNN models on emotion datasets, achieving strong performance. Du et al. [7] proposed a CNN-LSTM hybrid model with attention, improving classification accuracy.

Sadeh-Sharvit et al. [8] evaluated AI-based therapy systems and reported improvements in mental health treatment outcomes. Koh et al. [9] highlighted benefits and challenges of mobile mental health applications, including accessibility and privacy issues.

Conversational AI systems such as Wysa were studied by Beatty et al. [10], demonstrating scalability but lacking long-term validation. Transformer-based EEG models were further explored in [11] and [12].

C. Participatory Design and Chatbots (2021)

Danieli et al. [13] used participatory design methods to improve AI-based therapy systems by incorporating therapist feedback. Martinengo et al. [14] emphasized user-centered design in chatbot development.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

D. *Lightweight Models and Foundational Studies (2020–2019)*

Tursunov et al. [15] proposed lightweight CNN models for speech emotion recognition suitable for real-time use. Khanzada et al. [16] demonstrated CNN-based facial emotion recognition using the FER-2013 dataset.

Zhang et al. [17] and Khalil et al. [18] reviewed multimodal emotion recognition techniques and identified challenges in real-world deployment.

E. *Key Observations*

- Shift from unimodal to multimodal systems
- Increased use of transformer-based architectures
- Trade-off between accuracy and system complexity
- Challenges in privacy, engagement, and deployment

III. METHODOLOGY

A. *System Overview*

The proposed system is an AI-based mental health monitoring platform designed to provide real-time emotional analysis and adaptive support. Unlike earlier versions, the system follows a **multimodal and full-stack architecture**, integrating facial emotion recognition, voice tone analysis, and text sentiment evaluation.

The system continuously monitors user inputs and generates context-aware responses through a conversational AI module. It supports session tracking, journaling, and emotion analytics, enabling users to understand their emotional patterns over time.

B. *System Architecture*

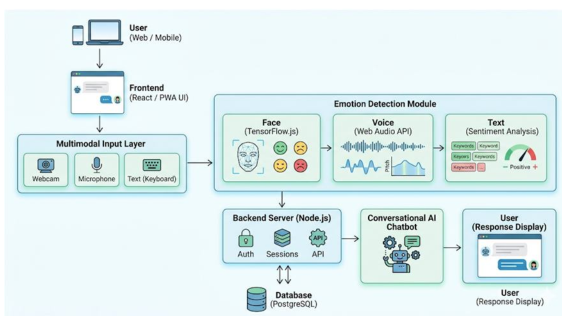

Figure 1: System Architecture of the Proposed System

Explanation:

Figure 1 illustrates the overall system architecture, which follows a client-server model. The frontend provides an interactive interface accessible via web and mobile browsers.

User inputs are collected through webcam, microphone, and text. These inputs are processed by multimodal emotion detection modules.

The backend handles authentication, session management, and API communication. Emotional data and session logs are stored in a PostgreSQL database. The chatbot generates real-time adaptive responses based on detected emotional states.

C. *Technology Stack*

The system is implemented using modern web and machine learning technologies to ensure scalability and real-time performance.

The frontend is developed as a web-based application with Progressive Web App (PWA) support, enabling cross-platform accessibility. Facial emotion recognition is implemented using TensorFlow.js and face-api.js, allowing real-time processing directly in the browser.

Voice analysis is performed using the Web Audio API, which extracts features such as pitch, intensity, and tone. Text sentiment analysis is handled using Natural Language Processing techniques on the backend.

The backend is built using Node.js and Express, which manage authentication, session handling, and API communication. Data is stored in a PostgreSQL database (Neon DB), ensuring persistence of user sessions and emotional records.

The system also supports hybrid mobile deployment using Capacitor.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

D. Multimodal Emotion Detection Pipeline

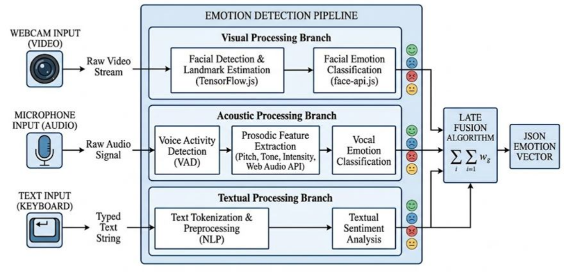

Figure 2: Multimodal Emotion Detection Pipeline

Figure 2: Multimodal Emotion Detection Pipeline

Explanation:

Figure 2 represents the workflow of the emotion detection system. The process begins with capturing user input through webcam, microphone, and text.

The input data undergoes preprocessing such as normalization and noise reduction. Feature extraction is then performed for each modality, including facial features, voice characteristics, and textual sentiment.

These features are analyzed to classify emotions such as happy, sad, angry, or neutral. The system continuously updates emotional states in real time, ensuring dynamic and accurate monitoring.

E. Conversational AI and Response Generation

After detecting the user's emotional state, the system generates adaptive responses using a conversational AI module. The chatbot provides supportive and context-aware replies, helping users manage their emotional well-being.

The responses are dynamically adjusted based on the detected emotion, ensuring personalized and empathetic interaction.

F. Session Management and Analytics

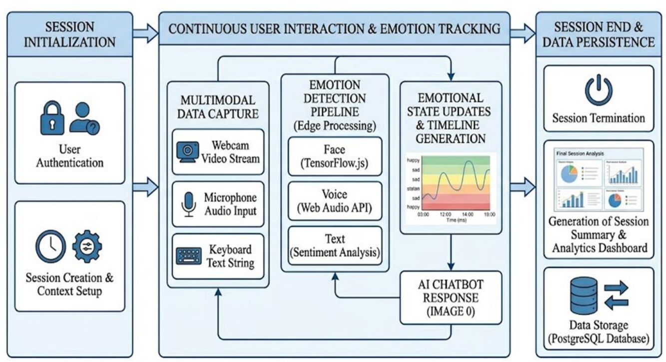

Figure 3: Session Flow and Emotion Tracking

Figure 3: Session Flow and Emotion Tracking

Explanation:

Figure 3 illustrates the session workflow. A session begins when the user initiates interaction with the system. During the session, emotions are continuously detected and stored along with timestamps.

The system maintains an emotion timeline and generates insights through analytics dashboards. At the end of the session, all data is stored in the database for future analysis and visualization.

G. Mobile Accessibility

The system supports mobile access through Progressive Web App (PWA) functionality. It allows users to access features such as emotion detection and chatbot interaction directly from mobile browsers.

Capacitor integration enables hybrid mobile deployment, which is currently under development.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

IV. RESULTS AND DISCUSSION

The proposed multimodal mental health monitoring system was tested in real-time environments using webcam, microphone, and text inputs. The system demonstrated efficient performance with low latency and smooth user interaction across different devices.

The integration of facial emotion recognition, voice tone analysis, and text sentiment analysis improved the robustness and reliability of emotion detection compared to single-modality systems. The system was able to dynamically update emotional states and provide context-aware responses through the chatbot.

The conversational AI module successfully adapted its responses based on detected emotions, creating a more engaging and empathetic user experience. Additionally, session tracking and emotion analytics enabled users to monitor emotional patterns over time, contributing to improved self-awareness.

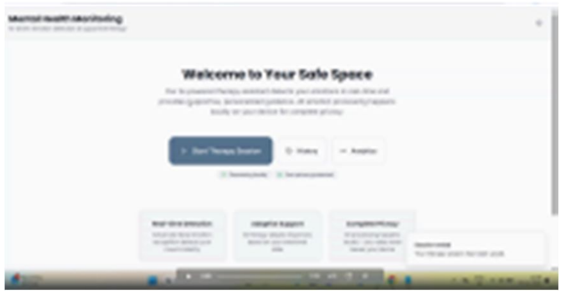

Figure 4: Home Interface

Explanation: Figure 4 shows the home interface of the system, which acts as the entry point for users. It provides options such as starting a session, viewing analytics, and accessing journaling features. The interface is designed to be simple and user-friendly.

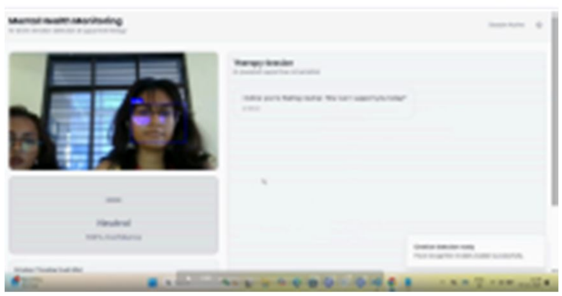

Figure 5: Emotion-Aware Chatbot Interface

Explanation: Figure 5 illustrates the chatbot interface integrated with real-time emotion detection. The chatbot interacts with the user while simultaneously analyzing facial expressions, voice tone, and text input. Based on the detected emotion, it generates adaptive and supportive responses.

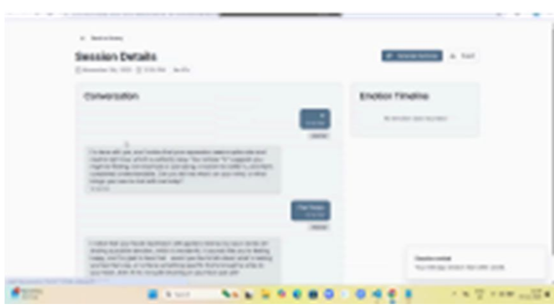

Figure 6: Session Detail Page

Explanation: Figure 6 presents the session detail page, where users can review past interactions. It includes detected emotions, timestamps, and chatbot responses. This helps users understand their emotional patterns and track their mental well-being.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

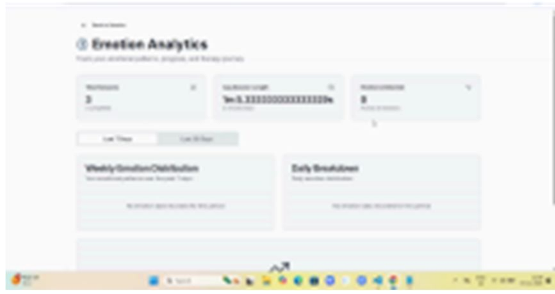

Figure 7: Emotion Analytics Dashboard

Explanation: Figure 7 shows the analytics dashboard, which visualizes emotional trends over multiple sessions. Graphical representations such as charts help users analyze fluctuations in emotions like happiness, sadness, anger, and stress.

V. CONCLUSION

This paper presents a multimodal AI-based mental health monitoring system that integrates facial emotion recognition, voice tone analysis, and text sentiment analysis to provide real-time emotional assessment and adaptive support. Unlike traditional approaches that rely on a single input modality, the proposed system combines multiple data sources to improve the accuracy and reliability of emotion detection.

The system has evolved from an initial prototype into a full-stack application, incorporating user authentication, session lifecycle management, journaling, and emotion analytics. The integration of a conversational AI chatbot enables context-aware and empathetic interactions, enhancing user engagement and overall experience.

Real-time processing capabilities and low latency ensure smooth interaction, while analytics features allow users to track emotional trends over time. The system demonstrates how AI can be effectively applied to create accessible, scalable, and user-friendly mental health support solutions.

Overall, the proposed system highlights the potential of integrating multimodal emotion detection with conversational AI to support early mental health awareness and intervention.

A. Future Work

Although the system demonstrates strong functionality, several enhancements can further improve its effectiveness:

- Integration of physiological signals such as heart rate and biometric data for deeper emotional analysis
- Development of personalized therapy models using machine learning and user behavior patterns
- Support for multiple languages and cultural adaptability to increase accessibility
- Enhancement of cloud infrastructure for secure and scalable long-term data storage
- Full deployment as a native mobile application for improved performance
- Integration with wearable devices for continuous and real-time health monitoring

VI. ACKNOWLEDGMENT

We would like to express our sincere gratitude to our project guide, Prof. Shraddha Katkar, for her valuable guidance, continuous support, and encouragement throughout the development of this project. Her expertise and insights significantly contributed to the successful completion of this work.

We also extend our thanks to the Department of Computer Science and Engineering, MIT Art, Design and Technology University, for providing the necessary resources and academic environment.

Finally, we acknowledge the use of open-source technologies such as TensorFlow.js, face-api.js, Node.js, and PostgreSQL, which played a crucial role in the development and implementation of this system.

REFERENCES

[1] J. Zhang, *et al.*, “EEG Multimodal Fusion Review,” *IEEE Access*, 2025.
[2] X. Zhu, *et al.*, “SOGPCN: Graph-Based Pseudo-3D CNN for EEG Emotion Recognition,” *Sensors*, 2024.
[3] Y. Pan, *et al.*, “DuA: Dual Attentive Transformer for Long-Term EEG Analysis,” *arXiv preprint*, 2024.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

[4] G. Xiao, *et al.*, "Four-Dimensional Attention Neural Network for Emotion Recognition," *Sensors*, 2024.
[5] Y. Chu, *et al.*, "MMESC: Multimodal Emotion Support Conversations," 2023.
[6] R. Yuvaraj, *et al.*, "3D Convolutional Neural Networks for SEED and SEED-IV Emotion Datasets," *Brain Sciences*, 2023.
[7] S. Sadeh-Sharvit, *et al.*, "AI-Based Behavioral Therapy Platform for Mental Health," *JMIR*, 2023.
[8] D. Danieli, *et al.*, "Therapist-Informed Participatory Design of AI Mental Health Agents," 2021.
[9] A. Khanzada, *et al.*, "Facial Emotion Recognition Using CNN on FER-2013 Dataset," in *Proc. IEEE Conf. Artificial Intelligence*, 2020.