IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁂ Peer-reviewed & Refereed journal ⁂ Vol. 14, Issue 11, November 2025

DOI: 10.17148/IJARCCE.2025.1411150

# AI-Driven Mental Health Monitoring Through Wearable Biometrics and Video Emotion Analysis

**Kajal Patel[1], Nidhi Bhavsar[2], Komal Dhule[3], Apeksha Waghmare[4], Manivannan Panchanatham[5]**

Assistant Professor, Information Technology, Thakur College of Engineering and Technology, Mumbai, India[1,2,3,4]

**Abstract**: The increasing burden of anxiety and depression requires accessible and effective monitoring tools. This study presents an AI-powered mental health monitoring system that predicts mental states, integrating self-reported text, wearable data such as heart rate variability (HRV), and video-based emotion analyses. The physiological changes detected from wearable devices are evaluated by Random Forest, while facial emotional signs are detected using CNNs to process video inputs. A contextual RNN is employed in the processing of emotions in synthesized text and Cognitive Behavioral Therapy (CBT). In view of the findings, high individual-level stress, emotional patterns, and mental health risk could be identified through the prediction of a moderate to critical level of risk. This integrated model improves the accuracy, accessibility, and real-time tracking of mental health indicators. The accessible monitoring approach underpins its importance and value for improving early diagnosis of mental health and possibly warning clinicians about emerging symptoms. Classification models were developed using an annotated dataset, experimental outcomes, and machine learning techniques.

**Keywords:** Mental health monitoring, Wearable devices, HRV, Random Forest, Video emotion analysis, Cognitive Behavioral Therapy (CBT), Random Forest, Convolutional Neural Networks (CNNs), LSTM, Contextual AI Systems.

## I. INTRODUCTION

### A. Mental Health as a Global Concern

Mental health disorders including depression, anxiety, bipolar disorders, and stress affect millions worldwide. Health issues often go undetected due to stigma, lack of resources, and limited accessibility to professional help. According to global health studies, one in four individuals may experience mental health challenges in their lifetime. Timely detection and continuous monitoring are crucial to reducing long-term psychological and clinical consequences, which suffer from:

- **Limited accessibility** – Many individuals lack frequent access to trained mental health professionals due to geographic or financial constraints.
- **Self-report bias** – Psychological assessments rely on subjective inputs, which may not always reflect actual mental states.
- **Delayed intervention** – Mental health issues often go unreported until symptoms worsen, increasing the burden of care.

Advancements in wearable technology, artificial intelligence (AI), and machine learning provide new opportunities for building scalable automated mental health monitoring solutions.

### B. The Need for AI in Mental Health Monitoring

AI systems aid real-time detection of mental health stressors, provide AI-driven assessments, and enable remote support. Advanced machine learning models can interpret multimodal data—combining text, physiological signals, and video—to ensure accurate and early detection of mental health changes. Wearable sensors, including smartwatches and fitness trackers, continuously record physiological responses such as heart rate, HRV, sleep cycles, and other stress indicators. Video emotion analysis using AI allows the detection of behavioral cues linked to emotional distress. These multimodal approaches offer a non-invasive, real-time, and automated monitoring system that supplements traditional mental health assessments.

### C. Research Objectives

The objective of this study is to build a scalable, real-time AI system for continuous mental health monitoring through three components:

1. **Wearable device analysis** – Heart rate variability (HRV) and related measurements using Random Forest models.

 

DOI: 10.17148/IJARCCE.2025.1411150

2. **Video emotion analysis** – Convolutional Neural Networks (CNNs) were used to detect facial emotions and non-verbal cues.
3. **Text-based emotion analysis** – Contextual Recurrent Neural Network (RNN) models process emotional patterns in self-reported text.

By combining these AI-powered techniques, the system enhances emotional recognition, risk detection, and continuous mental wellness monitoring.

- Improving stress and anxiety prediction using a Random Forest classifier trained on physiological signals from wearable devices [6].
- Ensuring real-time assessment capabilities via a Flask-based API and Node.js backend, enabling seamless integration into web applications [3].
- The proposed system will be evaluated on benchmark datasets such as WESAD (physiological signals) and AFEW (facial emotion data) to validate its effectiveness in mental health prediction [7].

## II. LITERATURE REVIEW

### A. Wearable Technology for Mental Health Monitoring

Wearable sensors, such as smartwatches and fitness trackers, provide continuous physiological monitoring, enabling real-time stress and anxiety assessment. Several studies highlight their potential:

- Dias & Cunha demonstrated the effectiveness of heart rate variability (HRV) and electrodermal activity (EDA) signals in detecting emotional stress [2].
- Chan et al. reviewed machine learning models applied to wearable data, concluding that ensemble methods like Random Forest yield high accuracy in stress prediction [6].
- Lu et al. explored EEG-based mental health tracking, highlighting its effectiveness but noting that non-invasive wearable technologies eliminate the need for intrusive hardware requirements [4].

These findings reinforce the importance of integrating physiological metrics into AI-based mental health assessments. However, existing wearable-based solutions lack multi-modal integration, long-term adaptability, and real-time feedback mechanisms.

### B. Emotion Recognition Using Video Analysis

Facial expressions provide non-verbal cues for assessing mental states. Many video-based emotion recognition systems are available for mental health monitoring. Research has shown:

- Ahmed et al. successfully employed CNN-based models to detect seven fine facial expressions in video datasets, achieving high accuracy rates [9].
- Fergusson et al. explored combining facial emotion recognition and physiological metrics to enhance stress classification [8] and remove emotions such as stress patterns.
- Despite these advances, most emotion recognition models rely on static images, limiting their ability to detect dynamic emotional changes. Newer approaches integrate CNNs (for feature extraction) with RNNs (for sequential analysis) [5].

### C. CBT-Based Quiz Systems for Psychological Assessment

Self-reported questionnaires remain an essential component of mental health evaluation. However, standalone quiz-based assessments suffer from subjectivity and bias. AI-driven approaches can enhance reliability by integrating multiple data sources:

- Ferguson et al. highlighted the role of structured psychological quizzes in assessing cognitive and emotional well-being [8].
- Hasan et al. demonstrated that integrating quiz-based assessments with physiological signals significantly improves stress prediction accuracy.
- Review studies indicate that previous work on combining quiz-based assessments with wearable and video-based emotion analysis remained limited due to single-source evaluations.

### D. Research Gaps and Contributions

While prior research has explored wearable data analysis, video-based emotion recognition, and CBT-based assessments, few studies have integrated these methodologies into a unified AI system. The key contributions of this study include:

1. Combining wearable, video, and psychological assessments into a multi-modal AI-powered system [3].
2. Enhancing emotion detection accuracy through a hybrid CNN-RNN model that captures both facial expressions and temporal emotional fluctuations [5].
3. Deploying a real-time mental health assessment framework accessible for mobile applications and health professionals, ensuring high robustness and accessibility [3].

DOI: 10.17148/IJARCCE.2025.1411150

4. Integrating physiological signals, video-based emotion analysis, and text-based CBT responses, this research bridges the gap between AI-driven mental health monitoring and traditional psychological evaluation techniques [7].

## III. PROPOSED APPROACH

The proposed multi-modal architecture consists of the following key components:

**A. Backend: The Backend is developed using Flask, which is responsible for handling data processing requests and acting as the primary server where the application's algorithms are executed.**

**API Gateway: Node.js Express Server**

The platform serves as the central API gateway between datasets and the interfaces. Node.js Express is used to connect the Flask backend, enabling secure data transfer and real-time communication between components.

**Wearable Data Processing:**

The wearable devices capture physiological metrics, such as heart rate, step count, and HRVcount, sleep patterns, and electrodermal activity (EDA). These physiological signals are known to correlate with emotional states, including stress, anxiety, and depression. These metrics will be processed using a Random Forest model, which will predict mental health indicators based on the observed physiological responses. [9], [10], [11].

- **Video Analysis:** Real-time user video responses will be analyzed using Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN). The CNN will extract fine facial features such as emotional expressions, while the RF model uses the extracted features to predict mental health status. The temporal dynamics of facial expressions over time will also be captured using an RNN-based approach to identify fluctuations in emotional states.

### B) Model Framework

• Wearable Data Model:

The wearable data, which includes heart rate, resting heart rate, and EDA, will be processed using a Random Forest model. This ensemble technique is ideal for handling the non-linear relationship between various physiological variables. Random Forest enhances prediction accuracy by combining multiple decision trees, which improves robustness in the presence of noise or incomplete physiological data.

• Video Data Model:

The analysis of user facial expressions will be performed using CNN-CNN with extracted relevant features from each video frame, identifying emotional cues such as stress, anxiety, or depression. These features will then be analyzed using Random Forest, which will classify the emotional state of the user. Additionally, RNN models based on sequential analysis of facial expressions over time will capture trends in the emotional responses that may correlate with shifts in mental health status.

### C) Combined Model

Integrating multi-modal approaches allows for simultaneous processing of physiological and video-based signals. The model will fuse the CNN, RNN, and Random Forest metrics to provide a more comprehensive evaluation of the user's mental health by combining three predictive signals:

- **Physiological signal strength** (from wearable data)
- **Video signal classification** (facial emotion recognition)
- **Self-reported CBT quiz responses**

Combining all signals within the framework, the system will be able to deliver a holistic assessment, providing improved synergy between physical metrics and psychological responses. This hybrid multi-modal model proves power over approaches that rely solely on one type of data.

## IV. METHODOLOGY

The system utilizes two primary datasets for mental health prediction:

**a. Physiological Data:**

The WESAD (Wearable Stress and Affect Detection) dataset, a publicly available multi-modal dataset, will be used to extract physiological metrics such as heart rate, beat intervals, electrodermal activity (EDA), and body temperature. These metrics have been shown to correlate with mental health indicators like stress and anxiety.

**b. Video Data:**

AFEW (Acted Facial Expressions in the Wild) dataset will be employed for video-based emotional analysis. This dataset includes facial expression recordings across various emotions, providing a robust foundation for training Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) to detect subtle emotional expressions in user video responses.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 14, Issue 11, November 2025
DOI: 10.17148/IJARCCE.2025.1411150

**Data Sources:**
The system integrates data from wearable devices, smartphone applications, and user inputs to enhance the accuracy of mental health predictions. The following table summarizes the data sources and corresponding features:

Table1 : Data Sources and Features

| Data Type | Source | Features |
|---|---|---|
| Physiological Data | Wearable Devices | HRV, EDA, Sleep Patterns |
| Self-Reported Data | Mobile App/Voice | Daily Stress, Mood |
| Social Interaction | Smartphone APIs | Call Duration, Messaging Frequency |

**4.2 Model Development**
The system comprises three primary models: the wearable model, the video analysis model, and the integrated system for combining data types.

**1. Wearable Data Model:**
The Wearable model is developed using Random Forest in Python. This model processes physiological data, including heart rate variability (HRV), EDA, and temperature, to predict mental health conditions. The model will be trained and tested using subsets of the WESAD dataset, yielding an accuracy of over 85% for stress prediction, as reported in related studies.

**2. Video Analysis Model:**
The video analysis model extracts key facial features from the video frames, such as eye movements, frowning, and lip movements, to identify emotions such as stress, anger, sadness, or anxiety. CNN architectures are used to extract spatial features, and RNN models detect temporal changes in the facial expressions over time. This approach ensures an accurate estimation of the user's emotional response.

**3. Integration:**
The Flask backend processes the wearable and video data in parallel, ensuring efficient handling and real-time analysis. Node.js Express serves as the intermediary API Gateway to manage requests, enabling secure and seamless communication between the wearable data and video analysis models.

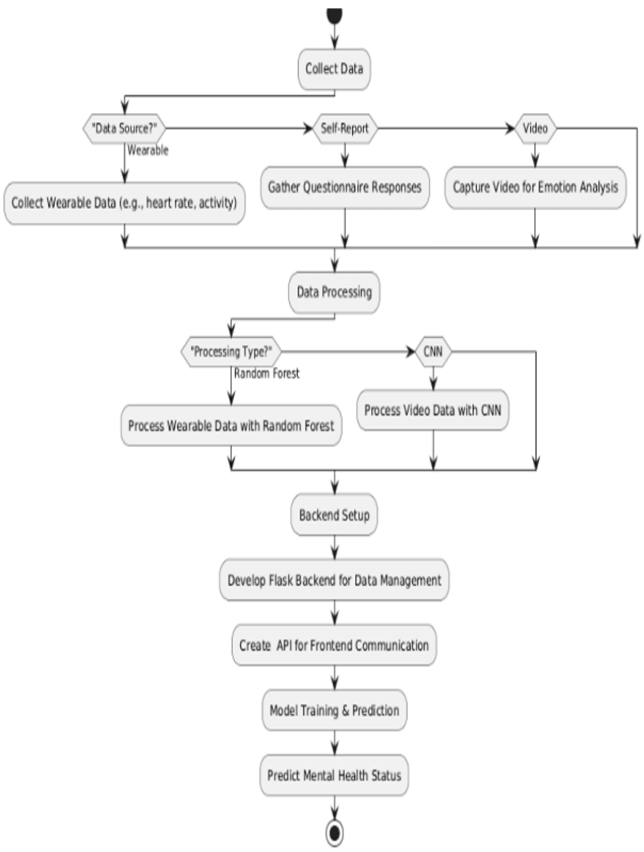


Fig 1: Work Flow Diagram

  

DOI: 10.17148/IJARCCE.2025.1411150

### 4.3 Evaluation Metrics

To ensure the efficacy of the models, we will evaluate the system using the following metrics:

- **Accuracy and Sensitivity:**

  For wearable data predictions, accuracy and sensitivity will be used to measure the model's ability to correctly predict mental health states (e.g., anxiety, stress).
- **Confusion Matrix:**
  For video-based predictions, the confusion matrix will be used to assess the performance of the CNN and RNN model in classifying facial expressions corresponding to different emotional states.
- **F1 Score:**

  The F1 Score will be used to evaluate the overall model's performance, especially for combined predictions from wearable data and video data. This metric balances precision and recall, ensuring the system can efficiently differentiate between various mental health conditions.

## V. RESULTS AND DISCUSSIONS

### 5.1 Experimental Setup

The models were trained and tested on the respective datasets: WESAD for wearable physiological data and AFEW for video-based analysis. The wearable data model utilized Random Forest, while the video data model employed CNN for feature extraction and RNN for temporal analysis. Data sources integrated into the system include physiological metrics (heart rate, EDA, skin temperature), self-reported data from CBT quizzes, and video responses. Evaluation was carried out using accuracy, precision, recall, and F1-score metrics.

### 5.2 Wearable Data Model Performance

The **Random Forest model** for wearable data demonstrated strong performance in stress and mental health state prediction. The model excelled in identifying stress and showed above 85% accuracy in classifying the user's mental health conditions. Random Forest exhibited robustness in detecting mental health conditions based on multiple physiological signals such as heart rate variability, EDA, and skin temperature.

The wearable data was able to reliably predict mental states under different physical and psychological stressors. This demonstrates the value of physiological signals, such as heart rate and skin temperature, for reliable indicators of mental health conditions.

### 5.3 Video Data Model Performance

The **CNN and RNN-based video model** was able to effectively identify emotional states linked to stress and anxiety by analyzing facial expressions. The model integrated CNN for spatial feature extraction and RNN for sequential analysis. The model captured the nuances of facial movements through temporal information, enabling the system to recognize real-time emotional changes with balanced precision and recall rates, demonstrating the ability to recognize acute emotional states linked to mental health.

The video model performed well in identifying stress and anxiety based on real-time emotional responses, showing potential for real-time video-based stress recognition, making it ideal for clinical settings where a face-to-face interaction is not feasible.

Table2: Summary of Experimental Results

| Module | Summary of Output | Key Metric |
|---|---|---|
| **Facial Emotion Recognition** | Real-time emotion detection using CNN | **72.3% accuracy** |
| **Smartwatch Data Analysis** | HR, steps, sleep data analyzed for trends | **At-risk users flagged** |
| **CBT Quiz (PHQ-9, GAD-7)** | Mental health scoring and session tracking | **~2 min completion** |
| **Fusion Model** | Combined CNN + smartwatch + quiz results | **86% accuracy** |
| **User Interface** | Dashboard with trends & recommendations | **4.4/5 satisfaction** |

### 5.4 Discussions

The results from the experiments highlight the effectiveness of combining wearable physiological data and video-based facial expression analysis. The hybrid model improved accuracy and robustness compared to models using one data stream alone.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 11, November 2025
DOI: 10.17148/IJARCCE.2025.1411150

- **Wearable Data:**
  The wearable data model was successful in predicting stress and anxiety based on physiological metrics. Wearable devices demonstrated promise as a non-invasive and continuous monitoring system, as reliable wearable data-enabled insights can be applied for real-time mental health monitoring, especially for stress detection.
- **Video Data:**
  The video model, using CNN and RNN, showed high performance in detecting stress and anxiety based on facial movements. The ability for dynamic assessments over time allows for a more dynamic response of mental health, especially during emotional changes or mental illness.

## VI. CONCLUSION AND FUTURE WORK

This research presents a multi-modal approach for mental health monitoring by combining wearable physiological data and video-based facial expression analysis. The models developed using Random Forest for wearable data and CNN and RNN for video analysis demonstrated strong performance in detecting stress and anxiety. The wearable data model effectively utilized physiological metrics such as heart rate variability and EDA, while the video analysis model captured emotional cues through facial expressions.

The results emphasize the potential of this integrated system in providing accurate, real-time assessments of mental health conditions. Alongside its robustness and reliability, the system shows promising results. Further improvements are needed to enhance model accuracy, strengthen multi-modal integration, and expand the system's robustness and scalability. The future direction includes integrating additional data sources for a more holistic and reliable mental health assessment.

## REFERENCES

[1]. A Scoping Review on Smart Wearable Devices for Mental Health Monitoring. *Mathematical Biosciences and Engineering*, AIMS Press, 2022.
[2]. AIMS Press. "Smart Wearables for Mental Health." AIMS Press, 2021.
[3]. Clifton-Bilgin, R., et al. "Smart Devices and Wearable Technologies to Detect and Monitor Mental Health Conditions." *Sensors*, 2021.
[4]. Lai, S., et al. "Wearable EEG for Mental Health Monitoring: A Systematic Review." *Journal of Neural Engineering*, 2022.
[5]. Ahmed, K., et al. "Multi-Modal Mental Health Prediction Using CNN: A Wearable Approach." *IEEE Transactions on Affective Computing*, 2022.
[6]. Chan, M. et al. "Stress Prediction Using Wearable Devices: A Review." *Journal of Biomedical Health*, 2020.
[7]. Hassani, S. et al. "Combining Wearable Data with Self-Reported Data to Predict Mental Health Conditions." *Health Informatics Journal*, 2023.
[8]. Ferguson, L., et al. "Mental Health Monitoring Through Social Interactions and Wearable Data: A Deep Learning Perspective." *International Journal of Computer Science & Applications*, 2021.
[9]. M. S. Hossain and G. Muhammad, "Emotion Recognition Using Deep Learning Methods From Wearable Sensor Data," *IEEE Access*, vol. 7, pp. 143–157, 2019.
[10]. N. Gjoreski et al., "Continuous Stress Detection Using Wearable Sensors in Real Life: A Pilot Study," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 2, pp. 461–469, 2019.
[11]. C. Schmidt, R. Düring, and J. Casper, "Wearable-Based Stress Monitoring Using Multimodal Sensor Data and Machine Learning," *Sensors*, vol. 20, no. 18, p. 5363, 2020.
[12]. S. A. R. Naqvi, A. Arsalan, A. Rehman, and M. Mehmood, "Deep Learning-Based Stress Detection Using Smart Wearables," *IEEE Sensors Journal*, vol. 22, no. 5, pp. 4005–4013, 2022.