© March 2026 | IJIRT | Volume 12 Issue 10 | ISSN: 2349-6002

# MindMate: A Personalized AI-Based Platform for Mental Health Support

Shriniwas Gavhane[1], Ayush Kalaskar[2], Vedant Deogadkar[3], Sahil Waghmode[4], Govind Pole[5]
[1,2,3,4,5]*Computer Science, Modern Education Society's Wadia College of Engineering, Pune, India*

*Abstract*—**Mental health disorders, including anxiety, depression, and chronic stress, have become a significant global concern, affecting individuals across all age groups and socio-economic backgrounds. The growing demand for accessible, affordable, and personalized mental health support has led to the emergence of artificial intelligence (AI)-driven solutions. This paper presents a comprehensive survey of recent advancements in AI-based mental health systems and introduces MindMate, a personalized, privacy-aware, multimodal AI platform designed to enhance emotional well-being.**

**Unlike traditional approaches that rely on a single source of data, MindMate integrates multiple modalities, including textual journaling, selfie-based facial analysis, and short video-based emotional assessment, to provide a holistic understanding of a user's mental state. The system performs sentiment analysis on journal entries, facial expression recognition from selfies, and behavioral and emotional cue extraction from video recordings. These inputs are processed using machine learning models and combined through a weighted multimodal fusion mechanism to generate accurate mood scores and meaningful psychological insights.**

**To ensure user trust and data security, the platform adopts a privacy-first architecture, where sensitive data such as images and videos are processed on-device using frameworks like TensorFlow.js and are not stored permanently. Only derived features such as emotion labels and mood scores are retained. Additionally, the system includes adaptive AI-driven interventions, such as personalized task recommendations, and early warning mechanisms for critical mental health conditions. In high-risk scenarios, the platform can trigger alerts to emergency contacts with user consent.**

**This survey emphasizes the effectiveness of multimodal AI in improving the accuracy, reliability, and personalization of mental health assessment systems. It also discusses current challenges, including model bias, data privacy concerns, and user engagement, while highlighting future research directions in intelligent mental health care systems.**

*Index Terms*— **Artificial Intelligence, Emotion Detection, Mental Health, Multimodal Fusion, Natural Language Processing (NLP), On-Device Machine Learning, Privacy-Preserving Systems, Sentiment Analysis, TensorFlow.js**

## I. INTRODUCTION

Mental health disorders, including anxiety, depression, and stress-related conditions, have emerged as a significant global public health challenge, affecting millions of individuals across diverse demographics. Despite increasing awareness, traditional mental health support systems often suffer from limitations such as restricted accessibility, high cost, social stigma, and lack of personalization. These challenges hinder early diagnosis and timely intervention, which are critical for effective mental health management [1], [2].

In recent years, advancements in Artificial Intelligence (AI), Machine Learning (ML), and Natural Language Processing (NLP) have facilitated the development of intelligent digital mental health platforms. These systems enable continuous monitoring, automated assessment, and scalable support, thereby improving accessibility and reducing dependency on conventional clinical settings [12], [15]. However, most existing solutions predominantly rely on unimodal approaches, such as text-based sentiment analysis or questionnaire-driven assessments, which fail to capture the multidimensional nature of human emotions [10], [19].

Human emotional states are inherently complex and are influenced by a combination of linguistic, facial, and behavioral cues. Relying on a single modality often results in incomplete or inaccurate assessments. Consequently, there is a growing need for multimodal mental health systems that integrate diverse data sources, including textual inputs, facial expressions, and video-based behavioral signals, to provide a more comprehensive and reliable understanding of an individual's psychological state [3], [4], [8].

In this context, this paper presents a survey of recent AI-based mental health monitoring systems and proposes MindMate, a privacy-aware, multimodal AI platform designed to deliver personalized mental health support. The proposed system integrates journaling-based sentiment analysis, selfie-based facial emotion recognition, and video-based behavioral analysis using deep learning techniques. By employing a weighted multimodal fusion strategy, MindMate generates dynamic mood scores, actionable insights, and early risk alerts. Furthermore, the system emphasizes a privacy-first architecture through on-device processing and secure handling of sensitive user data [5], [6], [11].

A. Key Contributions

1. Comprehensive literature review: Provides an overview of recent AI-driven mental health systems and identifies key limitations such as lack of multimodality and privacy concerns [1], [2], [13].
2. Multimodal framework design: Proposes an integrated system combining text, image, and video-based emotion analysis for improved accuracy [3], [4], [19].
3. Privacy-preserving architecture: Implements on-device machine learning to ensure sensitive user data is processed securely without storing raw media [6], [16].
4. Personalized support mechanism: Develops an adaptive system that generates mood insights, recommendations, and early risk alerts for proactive mental health care [8], [20].

## II. RELATED WORK

### A. Multimodal Feature Engineering

Recent studies highlight the importance of integrating multiple data modalities for accurate mental health assessment. Traditional approaches primarily relied on text-based sentiment analysis using Natural Language Processing (NLP) techniques to detect depression and anxiety from user-generated content [1], [2]. However, text alone cannot capture non-verbal emotional cues.

To overcome this limitation, researchers have incorporated facial expression recognition using Convolutional Neural Networks (CNNs), enabling detection of emotions such as happiness, sadness, and stress from visual data [3], [4]. Additionally, speech-based emotion recognition analyzes vocal attributes such as tone, pitch, and intensity to infer psychological states [17].

More advanced systems utilize multimodal feature fusion, combining textual, visual, and audio signals to improve prediction accuracy and robustness [8], [19]. Despite these advancements, challenges persist in synchronization, and noise handling.

### B. Deep Learning Architectures and Performance

Deep learning models have significantly improved the performance of mental health prediction systems. CNNs are widely used for image-based emotion recognition, while RNNs and Transformers are applied to sequential and textual data [5], [11].

Recent research explores transformer-based multimodal models and large language models (LLMs) for understanding complex emotional patterns and contextual interactions [11], [15]. These models demonstrate improved accuracy but often require large datasets and high computational resources.

Table I: Deep Learning Techniques in Mental Health Systems

| Technique | Application | Advantage | Limitation |
|---|---|---|---|
| CNN | Facial emotion | High accuracy | Needs large data |
| RNN / LSTM | Text & speech | Sequential learning | Gradient issues |
| Transformer / LLM | Context analysis | Long Dependencies | High cost |
| Multimodal Fusion | Emotion detection | Better accuracy | Complex integration |

### C. Conversational Systems and Safety Mechanisms

AI-based conversational agents and chatbots are increasingly used in mental health applications for providing real-time support and engagement. Studies show that AI chatbots can assist in emotional expression and early screening of mental health conditions [13], [14].

However, concerns regarding safety, reliability, and ethical considerations have been raised, especially in high-risk scenarios [20]. Recent works emphasize the need for fairness, accountability, and human-in-the-loop systems to ensure safe deployment of AI in mental health care [16].

Additionally, systems incorporating real-time alert mechanisms and intervention strategies have shown promise in preventing severe mental health crises [6], [9].

D. Gaps Addressed by MindMate

Despite significant advancements, existing systems exhibit several limitations, as summarized in Table II

Table II: Limitations of Existing Systems

| Limitation | Description |
|---|---|
| Lack of Personalization | Ge neric recommendations without user-specific adaptation |
| Limited Multimodal Usage | Reliance on single or partially integrated modalities |
| Privacy Concerns | Storage of sensitive user data such as images and videos |
| No Real-Time Intervention | Absence of proactive alert mechanisms for critical conditions |

## III. THEORETICAL FRAMEWORK

The proposed MindMate system is based on the principles of affective computing, multimodal machine learning, and privacy-preserving AI, enabling automated understanding of human emotions from multiple data sources. The framework integrates sentiment analysis, facial emotion recognition, multimodal fusion, and decision-making models to provide personalized mental health support.

A. Affective Computing and Emotion Recognition

Affective computing enables systems to detect and interpret human emotions using computational models. Facial emotion recognition techniques extract features such as facial landmarks, eye movement, and expressions, which are mapped to emotional states using deep learning models such as CNNs [3], [4].

Emotion classification can be represented as:

$$E = f(X_{face})$$

where $X_{face}$represents facial features and Edenotes the predicted emotion.

B. Natural Language Processing for Sentiment Analysis

The textual component uses Natural Language Processing (NLP) to analyze journal entries and extract emotional polarity. Sentiment analysis assigns a score indicating positive, negative, or neutral emotions [1], [2].

A commonly used sentiment scoring function is:

$$S = \sum_{i=1}^{n} w_i \cdot x_i$$

where $w_i$represents feature weights and $x_i$denotes textual features.

C. Multimodal Machine Learning and Fusion Strategy

To capture the multidimensional nature of emotions, the system integrates multiple modalities: text, selfie images, and video inputs. Each modality contributes differently to the final mood assessment.

The overall mood score is computed using a weighted fusion model:

$$M = w_1T + w_2I + w_3V + w_4Q$$

where:

- T= text sentiment score
- I= image (selfie) emotion score
- V= video emotion score
- Q= quiz score
- $w_1, w_2, w_3, w_4$= respective weights

This approach improves prediction accuracy and robustness compared to unimodal systems [8], [19].

Table III: Multimodal Data Contribution

| Modality | Feature Type | Contribution |
|---|---|---|
| Text | Sentiment | 40% |
| Image | Facial Expression | 30% |
| Video | Behavioral cues | 20% |
| Quiz | Self-reported data | 10% |

D. Deep Learning Models for Feature Extraction

Deep learning techniques are used for automatic feature extraction and classification. CNNs are applied for facial analysis, while transformer-based models process textual and multimodal data [5], [11]. These models learn hierarchical representations, improving emotion detection accuracy.

The prediction function can be generalized as:

$$y = f(X; \theta)$$

where Xis input data and θrepresents model parameters.

E. Privacy-Preserving Machine Learning

The system adopts a privacy-first approach by performing computations on-device using frameworks such as TensorFlow.js. Sensitive data such as images and videos are not stored; instead, only derived

features (emotion labels and scores) are retained [6], [16].
This can be represented as:

$$D_{stored} = f(X), \quad \text{where } X_{raw} \notin \text{storage}$$

F. Decision-Making and Risk Detection Model
The system incorporates a combination of rule-based logic and AI-driven analysis to detect critical mental health conditions. Risk assessment is performed by continuously monitoring the user's mood scores over time and comparing them against predefined threshold levels.
If the mood score consistently falls below a certain threshold or shows a declining trend over multiple observations, the system identifies it as a potential risk condition. In such cases, the platform initiates appropriate actions, such as generating alerts and recommending interventions. With user consent, notifications can also be sent to predefined emergency contacts to ensure timely support and early intervention [9], [20].

## IV. CORE TECHNOLOGY, MODELLING, AND PERFORMANCE

The MindMate system integrates advanced technologies from artificial intelligence, machine learning, and web-based frameworks to enable real-time, privacy-aware mental health monitoring. This section outlines the core technologies used, the system modelling approach, and performance considerations.
A. Core Technologies
The system is built using a combination of modern AI and web technologies:

- Frontend & Interface: Developed using web technologies to provide an interactive and user-friendly interface for journaling, video capture, and mood tracking.
- Authentication: Secure user authentication is implemented using Firebase Authentication or Clerk.
- AI Frameworks: TensorFlow.js is used for on-device machine learning, enabling real-time inference without sending sensitive data to servers.
- Natural Language Processing (NLP): Used for sentiment analysis of journal entries to detect emotional polarity and intensity [1], [2].
- Computer Vision: CNN-based models are used for facial emotion recognition from selfies and video frames [3], [4].
- Cloud Services: Firebase Cloud Messaging and Twilio APIs are used for notifications and emergency alerts.

These technologies collectively ensure scalability, responsiveness, and data privacy.

B. System Modelling
The system follows a multimodal data processing model, where inputs from different sources are processed independently and then combined.
Input Sources:

- Text (journal entries)
- Image (selfie-based emotion detection)
- Video (facial and behavioral cues)
- Quiz (self-reported psychological data)

Each input is processed through its respective model:

- NLP model → sentiment score
- CNN model → facial emotion score
- Video analysis → behavioral emotion score
- Quiz → numerical score

These outputs are integrated using a weighted fusion model, where different modalities contribute proportionally to the final mood score.

Table IV: System Modelling Components

| Component | Input Type | Output |
|---|---|---|
| NLP Module | Text | Sentiment Score |
| CNN Model | Image | Emotion Label |
| Video Analyzer | Video | Behavioral Score |
| Quiz Module | User Input | Numerical Score |
| Fusion Engine | All Inputs | Final Mood Score |

C. Performance Analysis
The performance of the system is evaluated based on accuracy, responsiveness, and usability:

- Accuracy: Multimodal fusion improves prediction accuracy compared to single-modality systems by capturing diverse emotional signals [8], [19].
- Real-Time-Processing: On-device inference using TensorFlow.js ensures low latency and immediate feedback to users.
- Privacy-Efficiency: Since raw media data is not stored and processing occurs locally, the system

reduces privacy risks and improves user trust [6], [16].
- Scalability: Cloud-based services enable efficient notification handling and system scalability for a large number of users.
- User-Engagement: Features such as mood tracking, personalized tasks, and weekly summaries enhance user interaction and consistency.

D. Comparative Performance

Table V: Comparison with Existing Systems

| Feature | Traditional Systems | MindMate |
|---|---|---|
| Multimodal Analysis | Limited | Yes |
| Real-Time Processing | Partial | Yes |
| Personalization | Low | High |
| Privacy Protection | Moderate | Strong |
| Emergency Alerts | Rare | Yes |

## V. DISCUSSION AND ANALYSIS

This section presents a critical analysis of existing AI-based mental health systems and evaluates the proposed MindMate framework in the context of current research trends, technological advancements, and practical challenges.

A. Comparative Analysis with Existing Systems

Existing mental health monitoring systems primarily rely on unimodal approaches, such as text-based sentiment analysis or questionnaire-driven assessments. While these methods provide useful insights, they often fail to capture the complexity of human emotions. Recent studies emphasize the effectiveness of multimodal systems, which integrate textual, visual, and behavioral data to improve prediction accuracy [3], [8], [19].

In comparison, MindMate adopts a fully multimodal approach by combining journal text, selfie-based facial analysis, video-based behavioral cues, and mood quizzes. Additionally, unlike many existing systems that rely on cloud-based processing, MindMate incorporates on-device machine learning, ensuring enhanced privacy and reduced latency [6], [16].

B. Key Trends in AI-Based Mental Health Systems

The literature reveals several emerging trends in the domain of AI-driven mental health:

- Shift towards Multimodal Learning: Integration of text, image, and audio data is becoming essential for accurate emotion recognition [4], [11].
- Adoption of Deep Learning and LLMs: Transformer-based models and large language models (LLMs) are increasingly used for contextual understanding and conversational support [14], [15].
- Focus on Privacy-Preserving AI: Techniques such as on-device learning and secure data handling are gaining importance due to ethical and regulatory concerns [16].
- Real-Time Monitoring and Intervention: Systems are evolving to provide continuous monitoring and proactive alerts for early mental health intervention [9], [20].

C. Strengths of the Proposed System

The MindMate platform offers several advantages over traditional systems:

- Multimodal-Integration: Combines multiple data sources to provide a comprehensive emotional assessment.
- Privacy-First-Design: Processes sensitive data on-device, avoiding storage of raw images and videos.
- Personalized-Recommendations: Adapts suggestions based on user-specific emotional patterns.
- Early-Risk-Detection: Identifies critical conditions and enables timely intervention through alerts.
- User-Engagement-Features: Includes journaling, mood tracking, and task-based activities to ensure consistent usage.

D. Challenges and Limitations

Despite its advantages, the proposed system faces certain challenges:

- Model-Accuracy-and-Bias: Emotion recognition models may produce biased or inaccurate results depending on training data [10], [25].
- User-Dependency: The effectiveness of the system relies on consistent user participation and data input.
- Computational-Constraints: On-device processing may be limited by device capabilities.

F. Practical Implications

The proposed approach demonstrates strong potential for real-world applications in:

- Personal Mental Health Monitoring: Enabling individuals to track and manage their emotional well-being.
- Preventive-Healthcare-Systems: Early detection of mental health issues before they become severe.
- Digital-Therapy-Support: Assisting therapists with data-driven insights and continuous monitoring.

## VI. CONCLUSION

This paper presented a comprehensive survey of recent advancements in artificial intelligence-based mental health monitoring systems, emphasizing the growing need for accessible, scalable, and personalized solutions in response to the increasing prevalence of mental health disorders worldwide. The analysis of existing approaches revealed that, although significant progress has been made in leveraging AI for emotional assessment, most systems remain constrained by unimodal data processing, and critical concerns related to data privacy and ethical deployment.

To overcome these limitations, the paper introduced MindMate, a privacy-aware, multimodal AI platform designed to provide continuous and personalized mental health support. The proposed system integrates multiple data modalities, including textual sentiment analysis, selfie-based facial emotion recognition, and video-based analysis, to capture a comprehensive representation of a user's emotional state. By employing a weighted multimodal fusion strategy, the system enhances accuracy, and contextual understanding of mood prediction compared to single-modality approaches.

A key contribution of the proposed framework lies in its privacy-first design, which leverages on-device machine learning techniques to ensure that sensitive user data, such as images and videos, are processed locally and not stored or transmitted unnecessarily. This approach not only addresses major privacy concerns but also improves user trust and system adoption. Additionally, the incorporation of adaptive recommendation mechanisms, and intelligent risk detection enables proactive intervention, allowing users to receive timely support and, in critical cases.

The findings of this study demonstrate that the integration of affective computing, deep learning, and multimodal data fusion can significantly enhance the effectiveness and reliability of digital mental health platforms. Furthermore, the proposed system highlights the importance of designing human-centered AI solutions that balance technological advancement with ethical considerations, fairness, and user privacy.

Despite its advantages, the proposed approach also presents certain challenges, including dependency on user participation, potential bias in emotion recognition models, and the need for extensive real-world validation. Addressing these challenges is essential for ensuring the reliability and clinical applicability of such systems.

In conclusion, this work underscores the transformative potential of multimodal AI in advancing mental health care and provides a foundation for future research in intelligent, privacy-preserving, and personalized mental health support systems. Future directions may include the integration of wearable sensor data, enhancement of real-time emotion detection capabilities, incorporation of explainable AI techniques, and collaboration with clinical professionals to validate and deploy such systems in real-world healthcare environments.

## REFERENCES

[1] S. Yeasmin, "Artificial intelligence for mental health monitoring: An umbrella review," *PMC*, 2025.

[2] M. Ali, "Artificial intelligence for mental health: A narrative review," *SAGE Journals*, 2025.

[3] W. Mostert, "Multi-modal emotion detection and tracking system using AI techniques," *Computers*, vol. 14, no. 10, 2025.

[4] G. E. Arzu *et al.*, "Adaptive multimodal emotion detection for mental health monitoring using deep learning," *Information Sciences*, 2026.

[5] R. Malik, "Emotion recognition for mental health assessment using transformer-based multimodal AI," *Research Square*, 2026.

[6] "AI4MentalWellness: Democratizing access to mood prediction," *International Journal of Science and Research Archive*, 2026.

[7] "AI-based multimodal depression detection framework," *JAAFR Journal*, 2026.

[8] "AI-driven multimodal mental health monitoring system," *IIETA Journal*, 2025.

[9] "AI-based emotion recognition in tele-therapy," *International Journal of Indian Psychology*, 2025.

[10] K. Hegde and H. Jayalath, "Emotions in the loop: A survey of affective computing for emotional support," *arXiv preprint*, 2025.

[11] X. Zhao *et al*., "Multimodal LLM for depression detection integrating visual and audio understanding," *arXiv preprint*, 2025.

[12] Su *et al*., "Deep learning in mental health outcome research," *Translational Psychiatry*, 2020.

[13] S. Siddals *et al*., "Experiences of generative AI chatbots for mental health," *npj Mental Health Research*, 2024.

[14] V. Sorin *et al*., "Large language models and empathy: A systematic review," *Journal of Medical Internet Research*, 2024.

[15] J. Torous *et al*., "Large language models as mental health providers," *The Lancet Psychiatry*, 2025.

[16] J. Moore *et al*., "Fairness and accountability in AI mental health systems," in *Proc. ACM FAT Conf.*, 2025.

[17] AIIMS Research, "Depression detection using speech patterns," *Times of India*, 2026.

[18] MUHS Study, "AI-based emotional assessment using facial cues," *Journal of Rare Cardiovascular Diseases*, 2026.

[19] K. Nguyen *et al*., "Multimodal sentiment analysis and fusion techniques," *Multimedia Systems*, 2025.

[20] J. M. Santos *et al*., "Evaluating clinical safety of AI in mental health responses," 2025.

[21] Y. Li, H. Chen, and Z. Wang, "Multimodal deep learning for mental health assessment using text, audio, and visual data," *IEEE Access*, vol. 11, pp. 45872–45885, 2023.

[22] Sharma, P. Gupta, and R. Mehta, "Emotion recognition in mental health systems using deep convolutional neural networks," *Expert Systems with Applications*, vol. 212, 2023.

[23] J. Brown, L. Smith, and K. Johnson, "Privacy-preserving machine learning techniques for healthcare applications," *IEEE Trans. Artificial Intelligence*, vol. 5, no. 2, pp. 145–158, 2024.

[24] M. Patel and S. Verma, "Real-time mental health monitoring using multimodal AI and wearable devices," *Sensors*, vol. 24, no. 3, 2024.

[25] Kumar, N. Singh, and A. Kaur, "Explainable AI for mental health diagnosis: Challenges and opportunities," *Artificial Intelligence in Medicine*, vol. 137, 2025.