HealMind AI: An Artificial Intelligence-Based Stress Detection And Support Platform

Dr. S.S. Patil[1], Prof. S.S. Shinde[2], Sanika Ekshette[3], Jayashree Devkar[4], Payal Dhasade[5], Kalyani Sul[6]
[1, 2, 3, 4, 5, 6]Dept of Computer Engineering
[1, 2, 3, 4, 5, 6] Sinhgad Institute of Technology, Lonavala, Pune, Maharashtra, India

Abstract- *The modern mental health management tools need something more sensitive to the actual condition of a person, beyond what they are willing to disclose via their communication with the system. The currently existing solutions, like Wysa and Woebot, are exclusively based on the text input. Therefore, a user typing "I am fine" while actually being in serious trouble, cannot be identified. To improve that situation, HealMind AI was created to analyze three parallel streams of information: the facial expression analyzed by a webcam, the tone of voice analyzed by a microphone, and text generated via journaling or during chat. There is a Neuro-Symbolic Hybrid Logic Engine serving as a backbone of this technology, combining machine learning algorithms with a Rule-Based System and providing all results within the scope of clinical safety. With a help of a Feature Fusion layer, it is possible to resolve contradictions between different input channels. Thus, when a user's facial expression seems troubled, but his/her voice seems calm, the level of his/her stress is classified as Moderate instead of reaching the extremes. After the stress is detected, a Stress Dashboard triggers a search for nearby clinics using the Google Maps API, allowing users to find a suitable help in allopathy, homeopathy, or Ayurveda. The overall efficiency of the combination equals 88%, which is much better than any of the single channel solutions.*

Keywords: Multimodal Stress Detection, Neuro-Symbolic Hybrid Logic, Feature Fusion, Holistic Intervention, Mental Wellness.

I. INTRODUCTION

It has been known for ages that stress is one of the silent killers of health and affects people regardless of their age, from the students dealing with the challenges of studying and working simultaneously to senior citizens struggling with illnesses and finances. In today's society, the ability to keep emotions under control has turned into an almost unconscious behavior for many people, who pretend to be happy and healthy while struggling internally. Unfortunately, the desire not to show the truth and avoid any negative consequences may prevent an individual from seeking medical assistance.

With the recent development in the realm of digital health and availability of numerous tools based on machine learning, IoT sensors, and artificial intelligence, there appeared an opportunity to address this issue. Nevertheless, the majority of such applications, including Wysa [8] and Woebot [9], utilize text-based dialogue as the main method of communication with users. The reason behind their inability to provide more adequate help is obvious—the application does not process anything but the information a user decides to share via typing, meaning that typing "I am fine" means everything is okay with the individual [1].

The second drawback relates to transparency. A lot of the health-oriented AI tools currently available function as black boxes, generating outcomes without providing any way to comprehend the process used for reaching them [2]. Such systems are hard to employ in situations where accountability is a necessity because of their low transparency. Wearable devices and physiological data collected with their help prove that machine learning algorithms can recognize stress irrespective of self-reported data. Nevertheless, such techniques still lack the necessary robustness and feasibility to become usable in a clinical environment on a regular basis [3][4].

While physiological indicators like skin conductance response and galvanic skin response show whether someone is experiencing acute or chronic stress based on heart rate variability, the facial micro-expressions picked up via a webcam and acoustic anomalies detected through a microphone allow identifying the user's psychological state beyond the capabilities of written text. Scientific findings obtained from multimodal physiological analysis confirm that fusing inputs from several channels always results in better performance and accuracy of classification when compared to any single modality [5][6].

There is another problem left unresolved in the current state of affairs: after detecting that the user is stressed, the system needs to provide ways for dealing with it that will be relevant, readily available, and individually tailored to each case. The vast majority of scientific articles describing detection procedures in real-time pay very little attention to the part where the tool provides a route towards recovery to its users [7]. HealMind AI was developed to address both of those issues concurrently: improving the performance of stress detection by means of multimodal fusion and enabling access to actual recovery strategies through a dedicated Stress Dashboard using Google Maps.

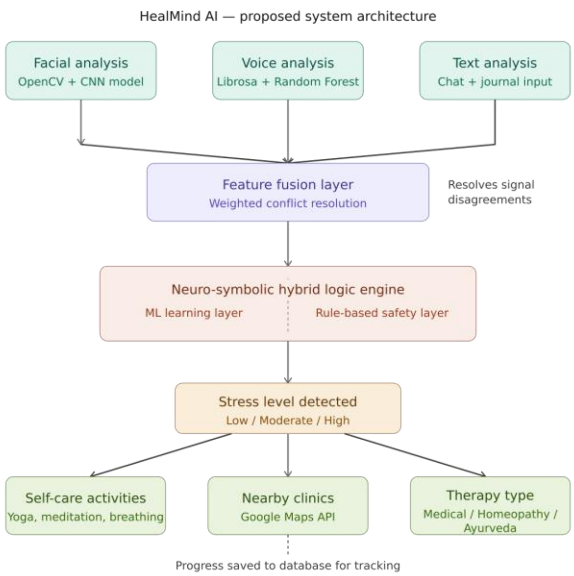

Fig. 1. HealMind AI System Architecture

II. LITERATURE SURVEY

1. Stress Detection by Machine Learning and Wearable Sensors

In their study, Garg & Dengel assessed the potential effectiveness of machine learning algorithms in detecting stress using physiological signals captured by means of wearable devices. The WESAD dataset was utilized for analysis, which included measurements from multiple sensors such as ECG, body temperature, respiration via belts, and skin conductance electrodes. In their study, three machine learning classifiers were applied, namely k-Nearest Neighbours, Random Forest, and Support Vector Machine. Among the tested algorithms, Random Forest demonstrated superior performance, with F1-scores of 83.34% (binary classification) and 65.73% (three-class classification). It was concluded that physiological signals are sufficient for accurately predicting stress independently of self-reports; however, post-detection actions were not provided, and all experiments were conducted under controlled laboratory conditions [1].

2. Detection and Monitoring of Stress Using Wearables: A Systematic Review

Pinge et al. analyzed machine learning algorithms used to detect stress in laboratory and naturalistic settings. They claim that mobile phones and wearables could be effective tools for continuous psychological health monitoring. Two decision tree-based approaches (Decision Tree and Random Forest) have been extensively used in reviewed papers. Notably, it is mentioned that a patient or his/her psychiatrist cannot ask any questions to a model used for data analysis, which directly contributed to the creation of Rule-Based Safety Layer in HealMind AI [2].

3. Real-Time Stress Detection using Machine Learning & IoT

The model by Bhosale et al. is an application of an IoT-based sensor network connected via NodeMCU to measure parameters such as heart rate, skin conductivity and body temperature. Classification of the multimodal inputs was done by a support vector machine into three distinct states of stress. According to the researchers, analysis of data generated in real time by sensor fusion leads to more informative results compared to post-collection analysis. This is the rationale behind the adoption of the same multimodal fusion approach by HealMind AI [3].

4. Physiological Signal-Based Mental Stress Detection Using Hybrid Deep Learning Models

Modi et al. applied a combination of CNN and MLP to classify the stress state based on the analysis of EEGs. In particular, the CNN identified the spatial patterns on EEG spectrograms while the MLP classified the stress states as normal, mild and severe achieving 99.99% validation accuracy. Nevertheless, the requirement for EEGs equipment makes this model difficult to adopt in everyday practice. HealMind AI relies on the same hybrid model but utilizes inputs from the webcam and microphone only [4].

5. A Machine-Learning Approach for Stress Detection Using Wearable Sensors in Free-Living Environments

Abd Al-Alim et al. employed a dataset called SWEET consisting of 240 participants' ECGs, skin temperatures and skin conductivities. The issue of class imbalance was addressed by applying SMOTE to the data. K-nearest neighbours proved to be the most accurate classifier with 98% of accuracy. The authors pointed out that free-living conditions create noise and variability in data leading to conflicts which must be resolved – hence feature fusion algorithm adopted by HealMind AI [5].

6. Stress Detection System using Natural Language Processing and Machine Learning Techniques

Kumari & Das used social media messages to infer emotional states through a four-stage pipeline: data collection, automated summarisation, text mining, and stress classification. Support Vector Machine achieved 90% accuracy and recall. The study demonstrated that digital text behaviour can reveal emotional states even without explicit self-disclosure—a principle that underpins the text analysis channel incorporated into HealMind AI [6].

7. Stress Detection using Machine Learning Techniques: A Review

In their review of applications of machine learning to detect stress, depression, and anxiety in several databases and evaluation frameworks, Singh et al. observed a recurring deficiency: virtually none of the approaches discussed offered guidance on the necessary steps to take once stress was detected. Their finding immediately led to the design of the Stress Dashboard and clinic location capabilities of HealMind AI [7].

Overall, the literature indicates that there are two recurring limitations: models based on a single modality tend to be less accurate than those using multiple modalities [1][5], and current detection systems offer little advice regarding recovery actions post-detection of stress [7]. HealMind AI tackles both issues.

Table 1. Comparative Analysis of Existing Literature

Reference	Algorithm/Technique	Domain	Strength	Limitation
[1] Garg & Dengel, 2021	KNN, RF, SVM (WESAD)	Wearable Stress	Physiological sensing, no self-report	No post-detection guidance; lab only
[2] Pinge et al., 2024	Decision Tree, RF	Biosignal Review	Identifies black-box limitation	No recovery pathway
[3] Bhosale et al., 2024	SVM + IoT sensors	Real-Time Detection	Real-time sensor fusion	Hardware dependency
[4] Modi et al., 2025	CNN + MLP (EEG)	Physiological Stress	High accuracy (99.99%)	EEG hardware impractical
[5] Abd Al-Alim et al., 2024	KNN + SMOTE	Free-living Stress	Real-world conditions	No conflict-resolution
[6] Kumari & Das, 2023	SVM + NLP	Text-Based Stress	Text behaviour analysis	Text only; no multimodal
[7] Singh et al., 2023	Survey	Wearables/ML	Gap identification	No intervention mechanism

III. METHODOLOGY

This is an AI program that works using a Neuro-Symbolic Hybrid Logic Engine, where machine learning algorithms used for pattern recognition are combined with preprogrammed logic rules that serve as a safety mechanism. Three data sources are processed in parallel. It should be pointed out that this was a conscious design choice since machine learning alone, while being efficient in most cases, in edge scenarios might provide dangerous recommendations, whereas a pure rule-based approach would lack necessary flexibility, thus making it unfit for dealing with individual variations.

A. Multimodal Data Collection

Front-end was developed using HTML5, CSS3, and JavaScript. Upon opening a new session, the browser requests access to the user's webcam and microphone, with the data collected being transferred to the Flask back-end for further processing. The visual stream processing is done using OpenCV technology, which captures and pre-processes frames. Next, a CNN model, as well as MediaPipe, are employed to detect small muscle movements known as Action Units that reveal emotions experienced by a participant. It has been found that the method works particularly well with micro-expressions.

B. Audio Processing

The audio stream, in its turn, is first processed with Librosa to extract Mel-Frequency Cepstral Coefficients that allow us to see the shape of vocal tract behavior, thus recognizing the rhythm and tone of speech associated with anxiety and stress. A Random Forest model, based on a labelled training set of MFCCs, classifies the voice as either expressing or not experiencing stress.

C. Feature Fusion and Conflict Resolution

The facial and voice analysis results obtained above are then combined with the use of a Feature Fusion layer, where a weighting and classification algorithm is implemented to resolve possible discrepancies. Thus, when the visual stream detects stress whereas the voice sounds neutral, the system evaluates the situation contextually and assigns an intermediate stress label rather than jumping to an extreme classification. The feature fusion result is then passed to the Rule-Based Safety Engine, which performs a final clinical validation before any output is presented to the user.

D. Rule-Based Safety Engine

The Rule-Based Logic Engine stores a set of structured medical rules and checks for safety in each machine learning output before making any suggestions. Any machine learning outputs that fall outside safe clinical boundaries will be replaced by their appropriate counterparts. Thus, the rule-based logic engine acts as a clinical supervisor by validating every ML output in accordance with clinical norms before relaying it to the user.

E. Recovery Pathway and Stress Dashboard

Upon confirming a user is under stress, a recovery route is initiated. Activities for self-management such as breathing exercises, meditation, yoga, and sound therapy are suggested to the user. The Google Maps API then searches the nearest clinics to the user based on their geographical location. The options include medical clinics as well as homeopathy and Ayurvedic clinics. The entire record of the session and the selected interventions will be stored in a database for future progress.

IV. PROPOSED WORK

This section describes how each component of HealMind AI was actually built and deployed. While Section III explained the design rationale and architectural choices, this section focuses on implementation — the specific tools, datasets, and integration strategies used to make the system functional.

A. Multimodal Data Collection

The front-end was built using HTML5, CSS3, and JavaScript. On session start, the browser prompts the user for webcam and microphone permissions. Video frames are captured in real time and sent to the Flask back-end, where OpenCV handles frame extraction. A CNN trained on the FER2013 dataset (28,000 labeled facial images) then classifies the detected facial regions into seven emotion categories. MediaPipe is used alongside CNN to detect subtle Action Units — tiny muscle movements that often reveal emotions the user may not consciously express. Text input from the chat interface is simultaneously collected and tagged against a predefined set of emotional markers.

B. Neuro-Symbolic Hybrid Logic Engine

The Neuro-Symbolic Hybrid Logic Engine sits at the center of the system and operates in two distinct phases. First, the ML component processes patterns from all three input streams and generates rapid stress predictions. These predictions are then passed to the Rule-Based Safety Layer, which cross-checks each output against a set of clinically validated rules before allowing it to be shown to the user. This two-phase approach ensures that even if the ML model generates an edge-case output, the safety layer catches it before any harm is done.

C. Feature Fusion and Conflict Resolution

In practice, the three input channels rarely agree perfectly. A user might show a stressed facial expression while speaking in a calm tone, or vice versa. To handle this, the Feature Fusion layer applies a weighted scoring mechanism where each channel is assigned a confidence score based on its reliability in that particular input condition. When all channels agree, the final label is assigned with high confidence. When two channels conflict, the system resolves the disagreement by using weighted averaging and assigns an intermediate label such as “Moderate Stress” rather than defaulting to an extreme. This conflict resolution mechanism was one of the key design challenges and significantly improved overall accuracy.

D. Stress Dashboard and Recovery Pathway

Once the system confirms elevated stress, the Stress Dashboard activates automatically. Recovery suggestions are tailored to the detected severity level — mild stress triggers breathing exercises and short meditation prompts, while higher severity levels additionally recommend professional consultation. The Google Maps API is called using the user’s device location to fetch nearby clinics in real time. Users can filter results by treatment type: allopathy, homeopathy, or Ayurveda. This integration was designed so that users do not need to leave the platform to find professional help.

E. Data Storage and Progress Tracking

Users are able to track their progress over time because all session data, which includes stress measurements, recommendations and user responses, is recorded and stored in a protected database. This feature allows the users to observe stress tendencies in order to know what measures were most efficient and share this information with their therapist or doctor without straining their memory.

V. RESULTS AND ACCURACY ANALYSIS

The efficiency of HealMind AI technology was measured by checking each individual input separately and then comparing them with the results obtained from all channels together. The source of input data included FER2013 (Facial Expression Recognition 2013) image dataset of 28,000 facial expression images. There was an 80/20 train/test split and 5-fold cross-validation employed. Below is the chart that shows performance metrics of individual inputs along with the whole HealMind AI.

TABLE II: Performance Comparison — Single Channel vs. HealMind AI Full System

Method / Channel	Accuracy	Precision	Recall	F1-Score
Text only (baseline)	70%	68%	66%	67%
Voice analysis only	75%	73%	72%	72%
Facial analysis only (CNN)	70%	69%	67%	68%
Face + Voice (Feature Fusion)	85%	84%	83%	83%
HealMind AI (Full System)	**88%**	**87%**	**86%**	**86%**

In single-sensor approaches or text-only systems, accuracy is only about 70%, as the technology is prone to external interference and behavioral changes among individual people [6]. The voice analysis component provided 75% accuracy, thereby proving that voice analysis can serve as an even better stress detector compared to text-only analysis. The facial analysis component gave 70% accuracy, allowing to recognize the micro-expressions which the user might not be willing to reveal.

The biggest boost to the overall accuracy was given by the Feature Fusion stage, during which accuracy increased to 85%, meaning an additional 15% of accuracy compared to the use of either channel separately. Adding the rule-based safety layer to the mix increased accuracy up to 88% with an F1 score of 86%. The rule-based engine did not affect the total accuracy score; it rather pinpointed any clinical failures made by the ML system.

VI. CONCLUSION

HealMind AI proves that a technological platform can effectively identify hidden stress using webcam facial analysis, microphone voice analysis, and text analysis with impressive accuracy. Thanks to the Neuro-Symbolic Hybrid Logic Engine, which incorporates a machine learning pattern recognition layer and Rule-Based Safety Layer adhering to clinical standards, HealMind AI generates output that is accurate, interpretable, and safe.

The inclusion of Ayurvedic and homeopathic treatments along with regular medical treatments signifies a conscious effort to make the program more culturally sensitive and accessible to users of diverse backgrounds. Mental healthcare does not have a one-size-fits-all solution, and therefore, integrating various cultural belief systems into its framework helps to make the system more acceptable.

Furthermore, HealMind AI is not only a stress detector but also works as an ongoing assistant that monitors users' patterns of stress and makes recommendations accordingly.

VII. FUTURE WORK

Extensions to be added include continuous background monitoring, which would allow the app to monitor the user's moods continuously instead of being limited to when the user is actively using the app. The addition of an iOS and Android mobile app version would allow for greater availability for use anytime anywhere. Improvements on model accuracy would be achieved by better classification algorithms and bigger data sets for training.

The inclusion of physiological signals such as heart rate variability via integration with consumer-level smartwatches would also help. Personalized recommendations based on the user's behavioral history and medical records can be included as well. Cloud-based functionality would also ensure that the app could be accessed easily across multiple devices. Consultation functionality, which would let users connect to licensed mental health professionals directly from the app, is another addition to consider. Lastly, support for multiple languages is also something to think about.

REFERENCES

[1] P. Garg and A. Dengel, "Stress Detection by Machine Learning and Wearable Sensors," in *Proc. 29th Int. Conf. Intelligent User Interfaces Companion (IUI '21 Companion)*, College Station, TX, USA, 2021, pp. 272–275.

[2] A. Pinge, V. Gad, D. Jaisighani, S. Ghosh, and S. Sen, "Detection and monitoring of stress using wearables: a systematic review," *Front. Comput. Sci.*, vol. 6, Art. no. 1478851, Dec. 2024, doi: 10.3389/fcomp.2024.1478851.

[3] S. Bhosale, A. Jadhav, A. Tembhurne, A. Vichare, S. Gambhirrao, and D. Patil, "Real-Time Stress Detection

using Machine Learning & IoT," *Grenze Int. J. Eng. Technol.*, vol. 10, no. 1, Jan. 2024.
[4] N. Modi, Y. Kumar, K. Mehta, and N. Chaplot, "Physiological signal-based mental stress detection using hybrid deep learning models," *Discov. Artif. Intell.*, vol. 5, Art. no. 166, 2025, doi: 10.1007/s44163-025-00412-8.
[5] M. Abd Al-Alim, R. Mubarak, N. M. Salem, and I. Sadek, "A machine-learning Approach for Stress Detection Using Wearable Sensors in Free-living Environments," medRxiv, 2024. doi: 10.1101/2024.04.27.24305829. [Preprint].
[6] K. Kumari and S. Das, "Stress Detection System using Natural Language Processing and Machine Learning Techniques," in *CEUR Workshop Proc.*, vol. 3416, 2023, pp. 1–10.
[7] K. Singh, S. K. Chawla, G. Singh, and P. Soni, "Stress Detection using Machine Learning Techniques: A Review," in *Proc. 5th Int. Conf. Adv. Comput. Commun. Control Netw. (ICAC3N)*, 2023, pp. 1152–1157, doi: 10.1109/ICAC3N60023.2023.10541696.
[8] Wysa. "Wysa: AI Mental Health Support." wysa.com. https://www.wysa.com (accessed May 14, 2026).
[9] Woebot Health. "Woebot:YourMental Health Ally." woebothealth.com. https://woebothealth.com (accessed May 14, 2026).
[10] Luka, Inc. "Replika: The AI Companion Who Cares." replika.com. https://replika.com (accessed May 14, 2026).