

AI-based Multimodal Stress Detection and Physiological Tracking Application

Jansi Rani¹, Arul Moneesh E², Dharaneesh R³ and Mahesh Veeraiah⁴

¹⁻⁴Information Technology, Sri Ramakrishna Engineering College, Coimbatore, India

Email: jansi.sankar@srec.ac.in, arulmoneesh.2105066@srec.ac.in, dharaneesh.2105026@srec.ac.in, mahesh.2105056@srec.ac.in

Abstract— This study presents StressEase, an AI-powered platform for real-time stress detection using three data modalities: facial expressions, speech sentiment, and smartwatch physiological signals. The system aims to offer reliable, non-invasive stress monitoring through machine learning and deep learning techniques. A Convolutional Neural Network (CNN) is employed for facial expression analysis, achieving 90% accuracy. For speech sentiment analysis, an SVM classifier with TF-IDF features is used, achieving 96% accuracy. Smartwatch data, including heart rate and sleep patterns retrieved using the Google Fit API, is processed by an XGBoost model with 92% accuracy. These outputs are fused through an ensemble strategy to determine the overall stress level. The system is deployed as both a web and mobile application, developed using React for the front end, Flask for the backend, and SQLite for the database, offering real-time insights and personalized relaxation recommendations. The ensemble model's combined predictions demonstrate the system's high reliability and potential for real-world applications in health, workplace wellness, and mental health monitoring.

Keywords— AI-powered stress detection, multimodal analysis, facial expression recognition, speech sentiment analysis, smartwatch data, real-time monitoring.

I. INTRODUCTION

Stress has become a common problem in today's fast-paced world, impacting people in a variety of settings, such as personal life, companies, and educational institutions. The conventional approaches to stress assessment, which frequently depend on clinical examinations and subjective surveys, are insufficient for giving prompt and precise information about a person's stress levels. These traditional methods have the potential to affect general well-being and productivity by causing misdiagnosis, postponed actions, and a lack of individualized support. This project introduces Stress Ease, a cutting-edge AI-powered platform that uses a multimodal approach to deliver real-time stress detection and physiological tracking. Stress Ease seeks to improve the precision and dependability of stress monitoring by combining speech sentiment analysis, facial expression analysis, and wristwatch physiological data, providing a complete answer to the problems with conventional assessment techniques. A complex facial recognition system driven by deep learning models, including Convolutional Neural Networks (CNNs), is the foundation of Stress Ease. The platform can precisely analyse face micro expressions thanks to this technology, even when seen from multiple perspectives and under diverse lighting situations Stress Ease maintains excellent accuracy and consistency, guaranteeing dependable stress detection in a variety of settings, in contrast to traditional face recognition systems that could have trouble with partial

visibility or occlusion. The platform uses sophisticated machine learning algorithms to integrate audio sentiment analysis in addition to facial analysis. Stress Ease can determine and identify stress levels based on speech patterns by analysing vocal semantic content. A more comprehensive assessment of a person's stress level is made possible by this multimodal approach, which offers insights that single-modality systems frequently overlook. Additionally, Stress Ease improves the accuracy of stress detection by utilizing physiological data from wearables, such as heart rate and sleep habits. The platform can offer a comprehensive picture of a person's stress levels by analysing various physiological signals, allowing for prompt interventions and tailored stress management advice. Stress Ease has a dynamic dashboard that shows real-time insights about stress levels, behavioural patterns, and productivity measures in order to guarantee user interest and foster self-awareness. Users are empowered to keep an eye on their wellbeing and make knowledgeable decisions regarding their mental health because of this transparency. Administrators can also monitor team well-being and spot patterns across departments, encouraging a supportive environment and proactive stress reduction. Stress Ease offers a scalable, cloud-native solution that is available to both individuals and enterprises, thereby addressing the shortcomings of current stress assessment techniques. Conventional systems frequently lack adaptation and flexibility, which makes them inappropriate for remote teams and varied work contexts. Stress Ease, on the other hand, is made to be effective and user-friendly, enabling customers to get immediate feedback on their stress levels without the need for laborious hardware or human input. Stress Ease improves the accuracy of stress detection and encourages a proactive approach to mental wellness by combining cutting-edge AI technologies with a multimodal approach. The platform is positioned as a useful resource for anyone looking to successfully manage stress in their daily lives due to its capacity to offer real-time information and tailored advice.

II. LITERATURE REVIEW

Several past works have explored the use of AI and IoT for stress detection. Studies have utilized ECG sensors, facial emotion analysis, sleep parameters, and speech patterns to detect stress. However, very few approaches integrate all these modalities into a unified system. StressEase addresses this gap by combining three powerful detection streams in one scalable application.

Kavita Pabreja, Anubhuti Singh, Rishabh Singh, Rishita Agnihotri, Shriam Kaushik, and Tanvi Malhotra [1]:

This study, focusing on Indian working professionals, aimed to predict stress levels using machine learning algorithms. By collecting data through questionnaires addressing eighteen factors affecting mental health, they identified key stress-triggering features such as workload, work-life balance, and financial concerns. Among the algorithms tested, the Support Vector Machine Regressor demonstrated superior performance, highlighting the potential of machine learning in proactive stress management. The research also suggests the need for personalized interventions based on identified stress factors, which could significantly enhance workplace well-being.

Ashutosh Singh, Khushdeep Singh, Amit Kumar, Abhishek Shrivastava, and Santosh Kumar [2]:

This study explored mental stress detection among university students using various machine learning algorithms, including Decision Trees, Random Forest, and Support Vector Machines (SVM). Their study, based on a questionnaire dataset from 843 students, found that SVM achieved the highest accuracy of 95% in distinguishing between stress and non-stress conditions, emphasizing its effectiveness in academic settings. Additionally, the researchers noted that certain demographic factors, such as age and academic performance, also played significant roles in predicting stress levels, suggesting that customized stress detection models could further improve accuracy.

Peter Neigel, Andrew Vargo, Benjamin Tag, and Koichi Kise [3]: This study investigated cyclical stress patterns in university students using wearable devices over a 28-month period. Their findings revealed population-wide increases in stress markers during exams, New Year's, and job hunting seasons. This study underscores the potential of wearables in unobtrusive, real-time stress monitoring, providing valuable insights into temporal stress patterns. The researchers also explored how incorporating additional contextual data, such as sleep patterns and physical activity levels, could improve the predictive power of stress monitoring systems.

Moein Razavi, Samira Ziyadidegan, Reza Jahromi, Saber Kazeminasab, Vahid Janfaza, Ahmadreza Mahmoudzadeh, Elaheh Baharlouei, and Farzan Sasangohar [4]:

This scoping review analyzed 98 peer-reviewed publications on machine learning and deep learning techniques for stress detection and monitoring. The study identified Support Vector Machine, Neural Network, and Random Forest models as consistently exhibiting superior accuracy and robustness. The review also highlighted the importance of physiological parameters, such as heart rate and skin response, as prevalent stress predictors.

Furthermore, the review emphasized the increasing integration of multimodal data (e.g., combining physiological and behavioral data) for enhanced stress detection models.

Gideon Vos, Kelly Trinh, Zoltan Sarnyai, and Mostafa Rahimi Azghadi [5]: This systematic literature review focused on stress monitoring using wearable devices and machine learning. It emphasized the challenges in developing generic stress measurement models due to individual variability and the lack of large, labeled datasets. The study called for further research into generalizable machine learning models and the development of substantial datasets for improved stress detection. Additionally, the review noted the growing interest in developing real-time stress monitoring applications, which could assist in mental health interventions by offering timely insights and personalized recommendations.

III. METHODOLOGY

A. System Architecture Overview

In order to identify stress levels in real time, the Stress Ease system integrates multimodal data sources utilizing a modular, web- and mobile-based architecture. It consists of three main layers: a SQLite database for storing and retrieving session-level stress results, a Flask-powered backend for processing and model execution, and a React-based frontend for user interaction. Through a web or an app interface, users can engage with the system by submitting brief video recordings for facial analysis, audio recordings for speech sentiment analysis, and consenting to the retrieval of real-time physiological data through integration with smartwatches via the Google Fit API[7].

Through specified RESTful API endpoints, the frontend and backend can communicate, allowing for real-time feedback and asynchronous processing. Machine learning models specific to each modality are executed by the backend. To maintain uniformity between modules, a distinct session ID is created for each user interaction. Every model publishes findings to a common database and runs independently on its own data type. Using a weighted ensemble approach, the outputs from the three models are combined to get the final stress categorization.

B. Facial Expression Analysis (CNN)

Convolutional Neural Network Architecture

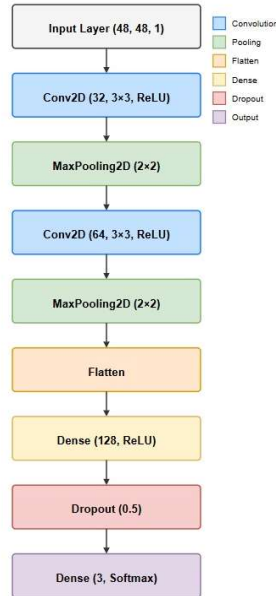


Fig. 1: CNN model architecture diagram for Facial

Micro expressions on the face are important markers of emotional stress. The 48x48 grayscale pictures used by the facial analysis module are taken from user-submitted video frames. A Convolutional Neural Network (CNN [6]) trained on a modified version of the FER dataset, which includes tagged emotional expressions, is fed the pre-processed frames. According to experimental results, the CNN [6] model demonstrated good generalization

performance over unknown data, achieving 90% training accuracy and 87% validation accuracy. A CNN [6]-LSTM architecture was also investigated in order to better capture temporal emotional patterns. Even though the LSTM component enhanced sequence modelling, in real-world implementation, the additional computational complexity did not considerably improve accuracy over the CNN alone [6]. As a result, the CNN [6] model was chosen due to its ability to balance computational efficiency and accuracy.

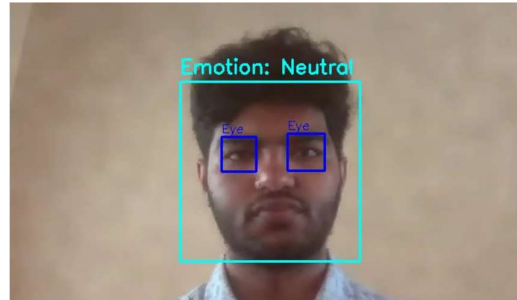


Fig 2: Facial stress Detection

C. Speech Sentiment Analysis (SVM + TF-IDF)

The speech-based stress detection module uses semantic information to assess the user's emotional tone. After users are asked to talk into the microphones on their devices, OpenAI's Whisper model is used to transcribe the audio. Term Frequency-Inverse Document Frequency (TF-IDF) is used to vectorize and preprocess the final text, highlighting emotionally charged terms. After a comparison analysis using Random Forest and Logistic Regression models, the Support Vector Machine (SVM) model was selected as the primary classifier. SVM and Logistic Regression both attained 96% accuracy, with Random Forest coming in second at 94%. SVM was chosen due to its great performance on moderately sized, high-dimensional text characteristics, stability, and reduced variance. Speech-derived stress levels are categorized by the model as Low, Neutral, or High.

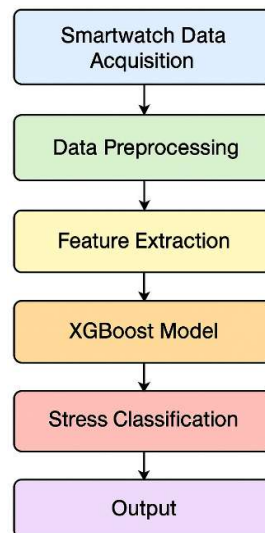


Fig. 3: Speech processing and SVM pipeline

D. Smartwatch-Based Physiological Monitoring (XGBoost)

Using wearable health measures, the physiological data stream offers insight into both chronic and acute stress. aspects including sleep duration, heart rate readings over the previous six hours (with 24 sampling intervals), and statistical descriptors like the minimum, maximum, average, and standard deviation of heart rate are among the aspects of the data, which is obtained from the Google Fit API [7]. Three machine learning algorithms—Random Forest, Support Vector Machine, and XGBoost—were evaluated in order to model these features [9]. Of these, XGBoost [9] outperformed Random Forest (90%) and SVM (87%), exhibiting the best accuracy at 92%. Its

ensemble-based methodology provided robustness against overfitting and improved handling of non-linear interactions. In order to assess physiological stress levels and evaluate smartwatch data, XGBoost [9] was selected for the final implementation.

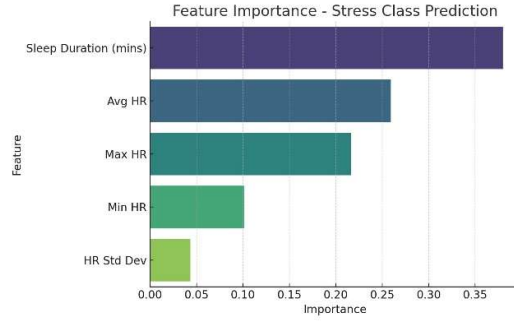


Fig. 4: Feature importance graph for Wearable data model

E. Ensemble Fusion Logic

The stress labels are categorized as Low, Neutral, or High is separately predicted by each of the three modules. A weighted ensemble fusion technique is used to combine these predictions into a single, cohesive stress classification. Weights are determined by the validation performance of each model: the speech model's score is 0.96, the smartwatch model's score is 0.92, and the facial expression model's score is 0.87. The weighted total of the model outputs is used to get the final stress score. The stress level is categorized as high if the total score is higher than a certain threshold. It is designated as Low if it drops below a lower threshold. Scores in the middle are regarded as neutral. When one of the input streams is noisy or unreliable, this method increases the system's robustness and enables it to balance between modalities.

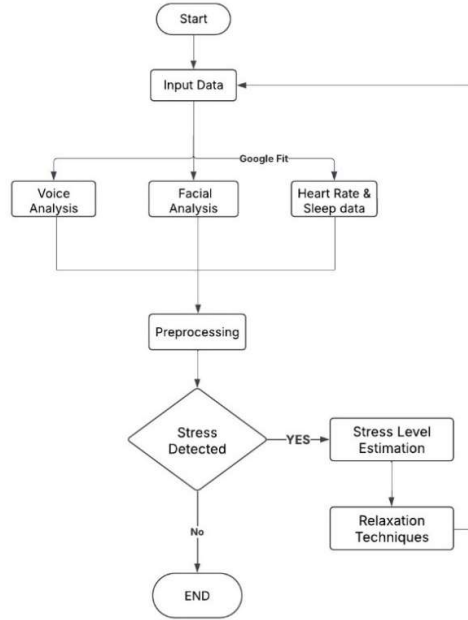


Fig. 4: Full system architecture diagram

F. Personalized Relaxation Recommendations

The technology dynamically suggests tailored wellness techniques targeted at reducing stress after identifying the general stress classification. These consist of white noise soundscapes, mindfulness and meditation activities, guided breathing sessions, and soothing music mixes. Based on the user's past stress patterns and the particular modality that signals elevated stress, recommendations are contextually modified. For instance, it might be

suggested that a user who experiences high levels of physiological stress on a regular basis follow physical relaxation exercises and sleep tracking. Users may react instantly to identified stress episodes thanks to these wellness interventions, which are shown through the user dashboard and accessible in real time.

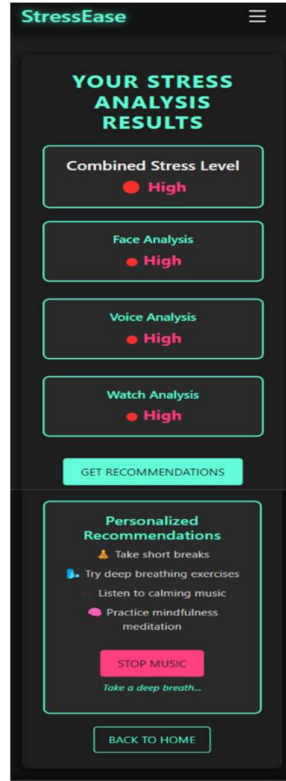


Fig 5: Recommendation Page

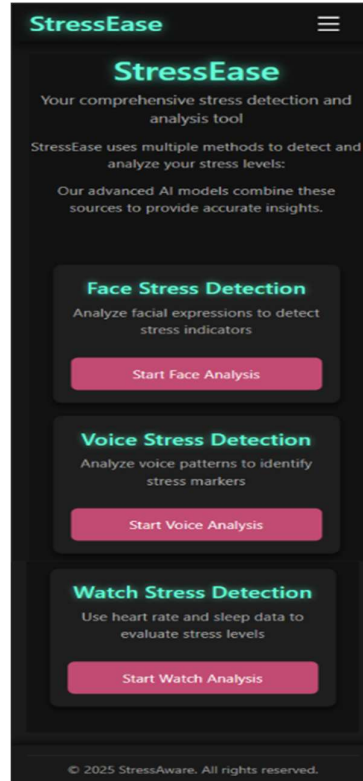


Fig. 6: Dashboard UI

IV. RESULTS AND EVALUATION

The assessment of separate AI models created for speech sentiment analysis, smartwatch physiological data analysis, and facial stress detection is presented in this part. Before being incorporated into the Stress Ease system, each model underwent separate training and validation to guarantee reliable feature extraction.

A. Facial Stress Detection Model Evaluation

A carefully selected dataset of stress-related facial micro expressions was used to train the Convolutional Neural Network (CNN) [6] model. With a 90% accuracy rate, it successfully captured minute variations in facial expressions and muscles that correspond to different stress levels. The efficacy of facial signals in stress detection tasks was validated by the model's high generalization across a variety of subjects.

B. Smartwatch Data Stress Detection Model Evaluation

Heart rate variability and sleep duration were among the physiological variables gathered from smartwatch sensors and analyzed using an XGBoost [9] classifier. With a 92% accuracy rate, this model demonstrated its capacity to decipher physiological patterns linked to stress reactions. The findings support the idea that wearable sensor data can be a trustworthy, non-invasive way to measure stress in real time.

C. Speech Stress Detection Model Evaluation

Stress levels were categorized using a Support Vector Machine (SVM) model that took into account speech characteristics such speaking rate, tone, and pitch change. With the highest individual accuracy of 96%, this model demonstrated how sensitive voice characteristics are to psychological and emotional stress. The performance of the SVM indicates that speech-based analysis, particularly when paired with additional data, is a potent medium for stress assessment.

V. CONCLUSION AND FUTURE SCOPE

With its non-invasive, real-time stress monitoring solution, the Stress Ease system showcases the effectiveness of AI-driven, multimodal approaches to stress detection and management. The technology offers a comprehensive tool for stress identification by combining physiological tracking via smartwatch data, speech sentiment evaluation, and facial expression analysis. Impressive accuracy has been achieved with the deployment of sophisticated machine learning models such as XGBoost [9], Support Vector Machines (SVMs), and Convolutional Neural Networks (CNN[6]s). The technology also provides tailored relaxation suggestions, including mindfulness exercises and guided breathing techniques, enabling users to take charge of their mental health. With its solid Flask backend, user-friendly React frontend, and real-time feedback features, Stress Ease is positioned as a scalable solution for enhancing mental wellness in a variety of contexts, including corporate wellness and personal health management. People may react to stress in real time thanks to the system's practical insights and suggestions, which promotes improved mental health. In terms of the future, Stress Ease has a wide range of possibilities for improvement. The incorporation of other biometric markers, such as skin conductivity and heart rate variability, is one important avenue that could improve the precision and accuracy of stress detection.

The capacity to monitor more precise physiological data will allow for even earlier identification and deeper insights into stress as wearable technology advances. Furthermore, adding contextual information about ambient elements like temperature, noise level, and illumination could enhance the system's comprehension of stressors. This would make stress management techniques even more individualized and efficient by enabling real-time modifications to suggestions based on the user's surroundings. By using reinforcement learning, the system may be able to adjust to changing stress patterns over time, guaranteeing ongoing enhancements and customization. Furthermore, Stress Ease has the capacity to grow considerably. When the system is used in bigger settings, such workplaces, colleges, or medical facilities, it may offer useful information on the general level of stress, which would be advantageous to communities and organizations.

Making the switch to a cloud-based infrastructure would enable a larger user base, making it an effective instrument for extensive monitoring of mental wellness. Lastly, partnerships with healthcare providers may increase the system's usefulness by allowing mental health practitioners to use it as an early diagnostic tool. By identifying those who may be at risk for stress-related illnesses, integration with clinical systems may enable early intervention and individualized treatment.

In summary, even though Stress Ease has already made great progress in AI-driven stress detection, there are still a lot of chances to improve its features and broaden its application. Stress Ease has the potential to revolutionize the way stress is monitored and controlled, promoting improved mental health globally, thanks to advancements in sensor technology, real-time flexibility, large-scale deployment, and healthcare integration.

REFERENCES

- [1] "Machine Learning Algorithms to Predict Stress Levels Among Working Professionals in India," K. Pabreja, A. Singh, R. Singh, R. Agnihotri, S. Kaushik, T. Malhotra, International Journal of Advanced Computer Science and Applications, Volume 14, Issue 5, pp. 232-241, 2023.
- [2] "Mental Stress Detection Among University Students Using Machine Learning Techniques," A. Singh, K. Singh, A. Kumar, A. Shrivastava, S. Kumar, Journal of Healthcare Engineering, Volume 2022, Article 9874521, pp. 1-12, 2022.
- [3] "Cyclical Stress Patterns in University Students: A Long-term Study Using Wearable Devices," P. Neigel, A. Vargo, B. Tag, K. Kise, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Volume 6, Issue 3, pp. 1-28, 2022.
- [4] "Machine Learning and Deep Learning Techniques for Stress Detection and Monitoring: A Scoping Review," M. Razavi, S. Ziyadidegan, R. Jahromi, S. Kazeminasab, V. Janfaza, A. Mahmoudzadeh, E. Baharlouei, F. Sasangohar, Computers in Biology and Medicine, Volume 142, Article 105226, 2022.
- [5] "Stress Monitoring Using Wearable Sensors and Machine Learning: A Systematic Literature Review," G. Vos, K. Trinh, Z. Saranyai, M. Rahimi Azghadi, IEEE Transactions on Affective Computing, Volume 13, Issue 4, pp. 1825-1841, 2022.
- [6] C. S. Vinoth Kumar, S. Karthick Raja, and R. Pavithra, "A Review on AI Based Multimodal Stress Detection and Mental Health Monitoring System," International Journal for Research in Applied Science and Engineering Technology (IJRASET), vol. 12, no. 4, pp. 1781-1785, 2024.
- [7] "Effortless Activity Tracking with Google Fit," R. Faheem, M. Asim, A. R. Javed, and M. Beg, Proceedings of the 2016 IEEE International Conference on Smart Computing (SMARTCOMP), 2016, pp. 1-8.
- [8] "FusionNet: An End-to-End Deep Neural Network for Multimodal Stress Detection Using Facial, Speech and Physiological Signals," M. Chen, J. Han, D. Li, X. Sun, IEEE Journal of Biomedical and Health Informatics, Volume 26, Issue 3, pp. 1255-1267, 2022.

- [9] Yu Tian, Yu-Feng Li, Zhi-Hua Zhou. "Speech Emotion Recognition Using Support Vector Machine," International Journal of Computer Applications, Vol. 1, No. 20, Feb 2010, pp. 6-9.
- [10] "A Deep Learning Approach for Stress Detection Based on Multimodal Bodily Expressions," J. Rodriguez-Ruiz, O. Mata-Hernandez, O. Martinez-Mozos, Frontiers in Computer Science, Volume 4, Article 835330, 2022.
- [11] "StressWatch: Real-time Stress Monitoring Using Multimodal Data from Commercial Smartwatches," Y. Lin, J. Chen, Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Volume 7, Issue 1, pp. 1-25, 2023.
- [12] "Detecting Work Stress in Offices by Combining Unobtrusive Sensors: A Multimodal Approach," S. Koldijk, M.A. Neerinx, W. Kraaij, IEEE Transactions on Affective Computing, Volume 13, Issue 2, pp. 768-782, 2022.