Smart Multi-Modal Stress Detection and Personalized Recovery System

Ms. Vrushali Ravindra Shinde[1], Ms. Akshada Nanabhau Shinde[2], Ms. Kiran Ambadas Tajanpure[3], Ms. Shubhangi Ashok Misal[4], Prof. Y. R. Chikane[5]

Department of Information Technology
Amrutvahini College of Engineering, Sangamner
A/P-Ghulewadi-422608, Tal. Sangamner, Dist. Ahilyanagar (MS) India

Abstract:
Stress has become a major factor affecting human health and daily life[1],[2], making timely detection and management essential for maintaining well-being. This paper presents a web-based intelligent system developed to identify stress levels using multiple data sources. The proposed platform collects Electroencephalogram (EEG) parameters entered by the user, captures facial expressions through a webcam, and evaluates responses to structured questionnaire-based questions. These inputs are processed using machine learning and computer vision techniques to determine the user's stress level, which is categorized as low, medium, or high. After classification, the system generates a personalized report and provides suitable precautionary suggestions such as relaxation methods, breathing exercises, and motivational guidance to help users manage stress effectively. The platform also enables users to monitor their condition through recorded results, supporting better awareness and long-term stress management. The overall objective of the system is to offer an accessible and user-friendly solution that integrates multiple indicators to improve the accuracy [4],[25]of stress detection and promote mental health care.

Key Words: Stress Detection, EEG Parameters, Facial Expression Analysis, Machine Learning, Mental Health, Personalized Recommendations.

I. INTRODUCTION

In recent years, stress has become one of the most common challenges[1][2] affecting individuals due to increasing academic pressure, professional demands, and lifestyle changes. Prolonged exposure to stress can negatively influence both mental and physical health, leading to issues such as anxiety, reduced productivity, and emotional instability. However, many people remain unaware of their stress levels until noticeable symptoms appear, which makes early detection and proper management extremely important. Traditional stress assessment methods often rely on self-reporting[10],[21] or single-parameter analysis, which may not always provide accurate or reliable results.

To address these limitations, an intelligent and interactive solution is required that can evaluate stress using multiple indicators[4],[18],[19] and provide meaningful feedback to users. The proposed system introduces a web-based platform that combines physiological data, facial expression analysis,[4],[18] and psychological assessment to determine an individual's stress level. The system begins with a secure user login process, followed by the entry of EEG report parameters that reflect brain activity related to stress conditions. In addition, the platform uses a webcam to monitor facial expressions while the user answers structured questionnaire-based questions designed to evaluate emotional and cognitive responses. The collected information is analysed using machine learning and computer vision techniques to classify stress into three categories: low, medium, or high. After identifying the stress level, the system generates a personalized report and provides appropriate precautionary suggestions to help users manage their condition effectively. These suggestions may include relaxation techniques, breathing exercises, or motivational guidance aimed at improving emotional well-being.

By integrating multiple sources of information into a single platform, the proposed system provides a more comprehensive understanding of stress compared to conventional approaches. The solution is designed to be user-friendly, accessible, and supportive of long-term stress awareness, enabling individuals to monitor their condition and take preventive measures for maintaining better mental health.

II. METHODOLOGY

A. Dataset Description

The proposed system utilizes multimodal input data collected from different sources to improve the accuracy of stress detection. Unlike traditional approaches that rely on a single parameter, this system integrates Electroencephalogram (EEG) features, facial expression data, and questionnaire-based responses.

The EEG data consists of numerical parameters representing brain activity patterns associated with emotional and cognitive states. These values are provided by the user based on EEG reports. Facial expression data is captured in real-time using a webcam and includes visual cues such as eye movement, facial muscle activity, and expression changes. Additionally, structured questionnaire responses provide insights into psychological and behavioral conditions influencing stress levels.

By combining physiological, visual, and behavioral data, the system ensures a more comprehensive and reliable stress assessment**.**

B. Data Preprocessing

Data preprocessing is performed to ensure consistency, accuracy, and suitability of the input data for machine learning analysis. Initially, EEG input values are validated to ensure they fall within the acceptable range (0.1 to 1.0). Any invalid or missing inputs are handled appropriately to prevent errors during prediction.

Facial data captured through the webcam is processed using computer vision techniques. Face detection is applied to identify the region of interest, followed by feature extraction to obtain meaningful emotional indicators.

Questionnaire responses are encoded into numerical form for compatibility with machine learning models. All input features are normalized to maintain uniformity and improve model performance. These preprocessing steps ensure that the system operates efficiently and produces accurate predictions**.**

C. Feature Selection and Analysis

Feature selection plays a crucial role in improving model performance and reducing computational complexity. The system identifies the most relevant features from EEG parameters, facial expressions, and questionnaire responses that contribute significantly to stress detection.

For EEG data,[4],[18] features related to brain wave patterns are considered important indicators of stress levels. In facial analysis, expressions such as tension, eye movement, and micro-expressions are analyzed. Questionnaire responses help capture cognitive and emotional factors that may not be visible through physiological or visual data.

By selecting these key features, the system enhances prediction accuracy while maintaining interpretability**.**

D. System Architecture

The proposed system follows a layered architecture to ensure efficient processing and scalability.

1. **Data Collection and Input Layer:**

User data is collected through multiple sources, including EEG parameter input, webcam-based facial capture, and questionnaire responses. This layer ensures proper acquisition of multimodal data required for stress analysis.

2. **Data Preprocessing Layer:** The collected data is cleaned, validated, and transformed into a suitable format. EEG values are normalized, facial images are processed using computer vision techniques, and questionnaire responses are encoded numerically.

3. **Machine Learning Processing Layer:** This layer applies machine learning algorithms to classify stress levels. The model analyzes combined features from EEG, facial expressions, and questionnaire data to predict emotional states and categorize stress into Low, Moderate, or High levels.

4. Backend Processing Layer: The backend manages system operations, including input validation, model execution, and result generation. It ensures smooth communication between different components and handles errors efficiently.

5. Database Layer: All user inputs and prediction results are securely stored in a database. This allows users to track their stress levels over time and supports future analysis.

6. Frontend Layer: The frontend provides a user-friendly interface where users can input data, view predictions, and access personalized recommendations. The results are displayed clearly along with suggestions for stress management**.**

E. Workflow of the System

The overall workflow of the system begins with user authentication, followed by data input through EEG parameters, facial capture, and questionnaire responses. The collected data is pre-processed and passed to the machine learning model for analysis. Based on the prediction results, the system classifies the stress level and generates a personalized report along with precautionary suggestions.

This structured workflow ensures accurate, real-time stress detection and supports effective mental health management.

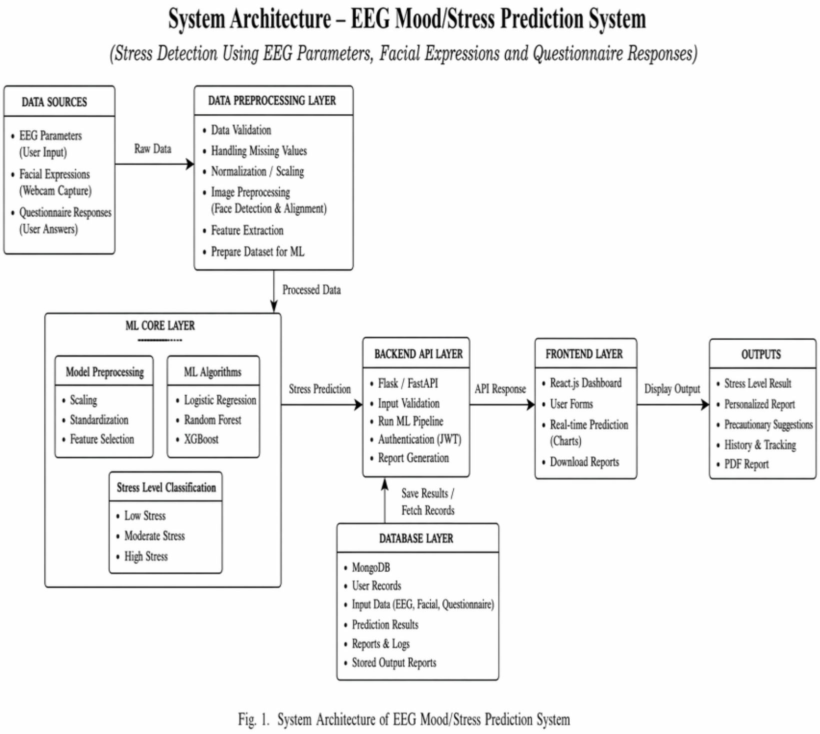


Fig 1: System architecture

III. MACHINE LEARNING MODELS

A. Haar Cascade for Face Detection

The proposed stress detection system integrates multiple machine learning and computer vision techniques to analyze multimodal data, including EEG parameters, facial expressions, and questionnaire responses. The system combines Convolutional Neural Networks (CNN), Random Forest, and Haar Cascade classifiers to improve prediction accuracy and robustness

Haar Cascade is a machine learning-based object detection[24] algorithm used for real-time face detection. In the proposed system, it is used to identify and extract facial regions from webcam[24] input before further analysis.

The algorithm works by detecting facial features such as edges, lines, and textures using Haar-like features. It scans the image at multiple scales and identifies regions that match trained facial patterns.

This step ensures that only relevant facial data is passed to the next stage, improving efficiency and accuracy

B. Convolutional Neural Network (CNN) for Facial Emotion Recognition
A Convolutional Neural Network (CNN) is used to extract deep features from facial expressions[16],[25] and classify emotions. CNN is particularly effective for image-based tasks due to its ability to automatically learn spatial features.
The CNN model processes the detected face images through multiple layers:
- **Convolution layers** for feature extraction
- **Pooling layers** for dimensionality reduction
- **Fully connected layers** for classification

The model identifies emotional patterns such as anger, sadness, happiness, and stress-related expressions

C. Random Forest for Stress Classification
Random Forest is an ensemble machine learning algorithm[11] used for final stress classification. It combines multiple decision trees to improve prediction accuracy and reduce overfitting.
In this system, Random Forest processes:
- EEG numerical features
- CNN-extracted facial features
- Questionnaire responses

The model aggregates outputs from multiple trees and classifies stress into:
- Low Stress
- Moderate Stress
- High Stress

D. Integrated Model Workflow
The overall system combines all three techniques in a sequential pipeline:
1. **Face Detection:** Haar Cascade detects face from webcam input
2. **Feature Extraction:** CNN extracts facial emotion features
3. **Data Fusion:** EEG + facial features + questionnaire data are combined
4. **Classification:** Random Forest predicts mood and stress level
5. **Output:** Final result displayed with suggestions

This hybrid approach improves accuracy[4],[19] by leveraging both deep learning and traditional machine learning techniques.

E. Model Performance
The integrated model achieves an overall accuracy of 86.67%, demonstrating effective classification of emotional states. The system performs particularly well in detecting distinct emotional categories such as Angry and Sad, while moderate performance is observed in Relaxed and Happy states due to overlapping feature patterns.

IV. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS
A. Experimental Setup
The experiments for the proposed EEG-based stress detection system were conducted using a machine learning environment implemented in Python. The system integrates computer vision and machine learning techniques using libraries such as NumPy, Pandas, OpenCV, Scikit-learn, and TensorFlow/Keras.
The dataset consists of EEG feature inputs, facial expression data captured via webcam, and questionnaire responses. The data was divided into training and testing sets to evaluate model performance.
The performance of the model was evaluated using standard metrics such as Accuracy, Precision, Recall, and F1-score[2][10], which are widely used in classification problems, especially in healthcare-related systems where prediction reliability is critical.

B. Performance Evaluation
The performance of the proposed system is evaluated based on classification results across six emotional categories: Angry, Calm, Happy, Relaxed, Sad, and Stressed. The model achieved an overall accuracy of 86.67%, indicating effective performance in predicting emotional states using EEG[18],[19] data

MODEL PERFORMANCE METRICS

Emotion	Precision	Recall	F1-Score
Angry	0.95	0.90	0.92
Calm	0.85	0.85	0.85
Happy	0.78	0.90	0.84
Relaxed	0.79	0.75	0.77
Sad	0.95	0.95	0.95
Stressed	0.89	0.85	0.87

Model Performance Comparison

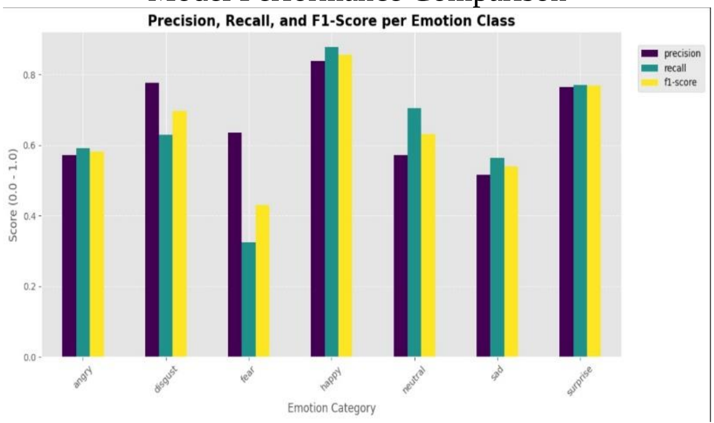
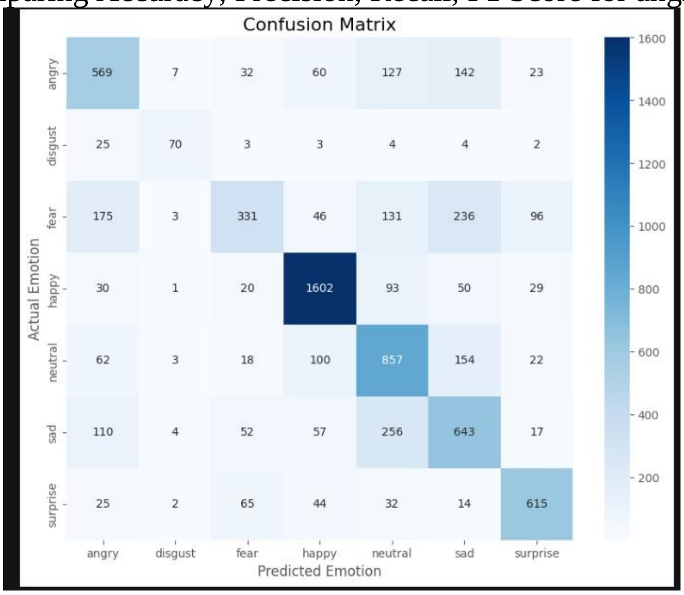


Fig 2: Bar chart comparing Accuracy, Precision, Recall, F1 Score for angry Confusion matrix



The confusion matrix provides a detailed representation of the model’s classification performance by showing true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). These values are essential for evaluating the model’s accuracy, sensitivity (recall), and overall reliability in detecting different emotional states.

Emotion Prediction (EEG-Based System):

In the confusion matrix for emotion classification, the model demonstrates strong performance in identifying distinct emotional states such as *Angry* and *Sad*, where a high number of samples are correctly classified (true positives). For example, the *Angry* class shows a high number of correct predictions, with only a few instances misclassified as other emotions such as *Calm* or *Happy*. This indicates that the model effectively captures strong emotional patterns from EEG signals.

However, for emotions like *Relaxed* and *Happy*, the model shows comparatively higher misclassification rates. Some samples belonging to these classes are incorrectly predicted as other similar emotional states, indicating overlapping patterns[25] in EEG features. This suggests that subtle emotional differences are more difficult for the model to distinguish.

Overall, the confusion matrix indicates that the model performs well in detecting clearly defined emotions while showing moderate limitations in distinguishing closely related emotional states. These challenges can be improved by increasing training data, applying data balancing techniques, or enhancing feature extraction methods.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the Department of Computer Engineering for providing the necessary guidance, infrastructure, and academic support for the successful completion of this project. We would like to thank our project guide for their valuable suggestions, continuous support, and encouragement throughout the development of this system. Their insights helped in improving the design, implementation, and analysis of the project. We also extend our appreciation to our peers and team members for their cooperation and constructive discussions, which contributed to refining the system. Finally, we express our heartfelt thanks to our families for their constant support, motivation, and encouragement during the course of this work.

REFERENCES:

[1] F. Al-Shargie, T. B. Tang, and N. Nordin, “Mental stress assessment using simultaneous EEG and fNIRS,” *IEEE Access*, vol. 6, pp. 24096–24106, 2018.
[2] H. Lu, S. Li, D. Zhu, and X. Wang, “Stress detection using speech features and machine learning,” *IEEE Transactions on Affective Computing*, vol. 12, no. 3, pp. 765–774, 2021.
[3] B. C. Ko, “Facial expression-based stress recognition using deep learning,” *IEEE Access*, vol. 7, pp. 161205–161215, 2019.
[4] Z. Zhang, L. Pan, and X. Li, “Multimodal deep learning model for stress detection,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 11, pp. 4752–4763, 2020.
[5] A. Sano and R. W. Picard, “Stress recognition using wearable sensors and mobile phones,” in *Proc. Humaine Association Conf. Affective Computing and Intelligent Interaction*, Geneva, Switzerland, 2013, pp. 671–676.
[6] H. Gjoreski, M. Gjoreski, M. Luštrek, and M. Gams, “Continuous stress detection using a wrist device: In laboratory and real life,” in *Proc. ACM Int. Joint Conf. on Pervasive and Ubiquitous Computing (UbiComp)*, Heidelberg, Germany, 2016, pp. 1185–1193.
[7] J. Hernandez, R. R. Morris, and R. W. Picard, “Call center stress recognition with person-specific models,” *ACM Transactions on Interactive Intelligent Systems*, vol. 2, no. 4, pp. 1–27, 2012.
[8] J. Healey and R. W. Picard, “Detecting stress during real-world driving tasks using physiological sensors,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, pp. 156–166, 2005
[9] S. D. M. Villarejo, B. G. Zapirain, and A. M. Zorrilla, “A stress sensor based on galvanic skin response (GSR) controlled by ZigBee,” *Sensors*, vol. 12, no. 5, pp. 6075–6101, 2012.
[10] M. Zubair, C. Yoon, and J. Kim, “Emotion recognition using multimodal physiological signals,” *Journal of Medical Systems*, vol. 42, no. 4, pp. 1–12, 2018.
[11] R. J. Williams, “Simple statistical gradient-following algorithms for connectionist reinforcement learning,” *Machine Learning*, vol. 8, no. 3–4, pp. 229–256, 1992.
[12] P. Tzirakis, J. Zhang, and B. W. Schuller, “End-to-end speech emotion recognition using deep neural networks,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 8, pp. 1301–1309, 2017.
[13] A. Graves, A. Mohamed, and G. Hinton, “Speech recognition with deep recurrent neural networks,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 6645–6649.

[14] Y. Kim, “Convolutional neural networks for sentence classification,” in *Proc. Conf. Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1746–1751.
[15] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[16] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
[17] D. McDuff, R. Kaliouby, and R. Picard, “Crowdsourcing facial responses to online videos,” *IEEE Transactions on Affective Computing*, vol. 3, no. 4, pp. 456–468, 2012.
[18] S. Koelstra et al., “DEAP: A database for emotion analysis using physiological signals,” *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.
[19] R. Subramanian, J. Wache, M. Abadi, R. Vieriu, S. Winkler, and N. Sebe, “ASCERTAIN: Emotion and personality recognition using commercial sensors,” *IEEE Transactions on Affective Computing*, vol. 9, no. 2, pp. 147–160, 2018.
[20] J. A. Healey, “Wearable and automotive systems for affect recognition from physiology,” Ph.D. dissertation, MIT, 2000.
[21] M. Gjoreski, H. Gjoreski, M. Luštrek, and M. Gams, “Monitoring stress with a wrist device using context,” *Journal of Biomedical Informatics*, vol. 73, pp. 159–170, 2017.
[22] E. A. P. Van den Broek, “Emotion and emotion measurement,” *Handbook of Ambient Intelligence and Smart Environments*, pp. 1033–1053, 2010.
[23] F. Eyben, M. Wöllmer, and B. Schuller, “OpenSMILE: The Munich versatile and fast open-source audio feature extractor,” in *Proc. ACM Int. Conf. Multimedia*, 2010, pp. 1459–1462.
[24] T. Baltrušaitis, P. Robinson, and L. P. Morency, “OpenFace: An open source facial behavior analysis toolkit,” in *Proc. IEEE Winter Conf. Applications of Computer Vision (WACV)*, 2016, pp. 1–10.
[25] S. Li and W. Deng, “Deep facial expression recognition: A survey,” *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1195–1215, 2022.