MENTAL HEALTH TRACKER

Mrs. Iswarya M M.E.,
Department of Computer Science and Engineering
(Assistant Professor)
Sri Sairam Engineering College
Chennai, India
Email:iswarya.cse@sairam.edu.in

Mrs. Susitra K M.E.,
Department of Computer Science and Engineering
(Assistant Professor)
Sri Sairam Engineering College
Chennai, India
Email:susitra.cse@sairam.edu.in

Harinarayanan A
Department of Computer Science and Engineering
Sri Sairam Engineering College
Chennai, India
Email:sec22cs109@sairamtap.edu.in

Lakshmi Narayanan S
Department of Computer Science and Engineering
Sri Sairam Engineering College
Chennai, India
Email:sec22cs073@sairamtap.edu.in

Elaiyavanan A
Department of Computer Science and Engineering
Sri Sairam Engineering College
Chennai, India
Email:sec22cs158@sairamtap.edu.in

2025 International Conference on Computing and Communication Technologies (ICCCT) | 979-8-3315-3757-9/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICCCT63501.2025.11020153

***Abstract*— Mental Health is a critical aspect of overall well-being, yet it often goes undiagnosed due to a lack of accessible tools for early detection. Our project presents a novel approach to mental health tracking using facial expression, voice recognition technologies and stress related words spoken by person. By analyzing key emotional cues through facial gestures and vocal tones, the system is designed to detect early signs of stress, anxiety, and depression. In addition to detection, the system offers personalized remedial measures aimed at relieving stress, such as guided breathing exercises, mindfulness prompts, and suggestions for professional mental health resources. The integration of real-time data analysis and machine learning models ensures that the system provides accurate and responsive feedback, promoting mental well-being in a non-intrusive, user-friendly manner. This paper outlines the architecture, implementation and potential impact of our mental health tracking system.**

Keywords*— *Mental health tracking, facial expression analysis, voice recognition, stress detection, machine learning, real-time data analysis, stress-relief interventions.

I. INTRODUCTION

Mental health plays a pivotal role in the overall well-being of individuals, influencing thoughts, behaviours and emotions. However, mental health disorders, such as stress, anxiety, and depression, often go undetected due to a lack of accessible, real-time tools for early diagnosis. According to the World Health Organization (WHO), over 264 million people globally suffer from depression, with many cases going untreated due to social stigma, unawareness, and the limited availability of mental health professionals. Traditional methods of diagnosing mental health issues often rely on self-reporting or clinical observation, both of which are subjective nd may not capture real-time emotional fluctuations. These limitations underscore the pressing need for accessible, automated tools that can monitor mental health continuously and provide timely intervention.

In this context, our project presents a novel solution by leveraging advanced facial expression and voice recognition technologies to develop a “Mental Health Tracker” system. The goal of our system is to detect early signs of emotional distress such as stress, anxiety, and depression-through non-intrusive, real-time monitoring of facial gestures and vocal tones. By combining these two input sources with advanced data processing techniques, the system provides a holistic view of individual’s emotional state.

Facial expressions are a powerful indicator of human emotions. Research shows that facial moments, such as raised eyebrows, frowns, and lip movements, can reflect emotions like sadness, anger, stress, and happiness. Similarly, voice tones convey emotional nuances through variation in pitch, rhythm, and intensity, which can signify mental states such as nervousness, calmness, or frustration. By analysing both facial and vocal data, this system can detect subtle emotional changes that might not be noticeable through traditional methods.

Beyond detection, our project emphasizes the importance of immediate and effective intervention. Upon identifying emotional distress, the system offers personalized, science-backed remedial measures. These interventions are designed to alleviate stress and include guided breathing exercises, mindfulness techniques, and recommendations for professional mental health support when necessary. The system is non-invasive, making it accessible for daily use without requiring significant changes in the user’s behaviour or lifestyle.

The core technology behind our system is driven by modern data analysis methods, enabling it to process emotional inputs quickly and accurately. The integration of real-time data processing ensures that the system can respond swiftly to fluctuations in the user’s emotional state, providing timely feedback and interventions.

In summary, our project addresses the critical gap in real-time, accessible mental health tracking by using technology to automate the detection of emotional distress. With the growing prevalence of mental health issues and the limited availability of professional care, this system has the potential to provide individuals proactive tools to manage their emotional well-being. This paper will explore the system’s architecture, technical implementation, and potential applications, highlighting how our project contributes to a more holistic approach to mental health monitoring.

II. LITERATURE SURVEY

The growing concern over mental health issues has driven researchers to explore innovative, technology-based solutions for real-time monitoring and early detection of psychological distress. Traditional methods of mental health diagnosis, such as clinical interviews and self-reported questionnaires, are often limited by their reliance on subjective reporting, which can lead to underreporting or misdiagnosis. This has spurred interest in the use of non-invasive, automated techniques like facial expression analysis and voice recognition, voice-based emotion detection, and integrated approaches to mental health tracking.

A. Facial Expression Recognition for Emotional Analysis

Facial expression recognition has been a well-researched area for detecting emotions such as happiness, sadness, anger, and stress. Early research in this field was largely inspired by the work of Paul Ekman [1], who categorized universal facial expressions that correlate with specific emotions. Advances in image processing techniques have enabled more accurate detection of these emotions through facial landmarks. Systems like “FaceReader” [2] utilize facial analysis to identify emotional states, with applications extending from mental health to customer experience and education.

Several studies have explored the use of facial expression recognition for mental health detection. For instance, a study by Liu et al. (2020) [3] presented a system that could recognize stress and anxiety by analyzing facial micro-expressions and eye movements in real-time. Their results demonstrated the effectiveness of facial cues as reliable indicators of emotional distress. Similarly, research by Zhang et al. (2019) [4] developed a stress-detection system that used facial expressions to predict mental fatigue and stress, demonstrating the feasibility of using facial data for continuous mental health monitoring.

Despite the promising results of facial expression recognition for emotion detection, it is important to note that emotions are multifaceted and can be influenced by various factors such as context and personality. Thus, while facial expression analysis provides valuable insights, it should be complemented by other indicators, such as voice data, to improve accuracy.

B. Voice-Based Emotion Detection

Voice, as a natural mode of communication, carries significant emotional information. Variations in pitch, tone, loudness, and speech rate can be indicative of a person’s emotional and mental state. Early work in speech analysis focused on speech recognition for linguistic purposes, but over the last decade, researchers have expanded into emotional speech analysis.

A seminal study by Schuller et al. (2013) [5] demonstrated that voice features could be used to detect emotions like anger, sadness, and happiness with considerable accuracy. Building on this, researchers have applied voice-based emotion detection in mental health applications, particularly in detecting stress and depression. A study by Cummins et al. (2018) [6] explored in clinical and non-clinical populations. Their system achieved promising results by using audio signal processing techniques to distinguish between depressed and non-depressed individuals.

More recently, studies have investigated voice analysis as a tool for continuous mental health monitoring. For example, a system developed by Yang et al. (2020) [7] applied voice recognition to detect mental fatigue in workers, showing that vocal changes could predict stress levels. Such systems have the potential to be integrated into everyday devices for real-time mental health assessment.

C. Integrated Approaches for Mental Health Tracking

While facial expression and voice recognition technologies have individually shown success in detecting emotions, combining these modalities yields a more comprehensive view of an individual’s mental state. Several studies have explored multimodal approaches that integrate facial and vocal data for more robust emotion detection.

An example of this is the work by Poria et al. (2017) [8], which developed a multimodal system combining facial expressions, voice tone, and text data to classify emotions across multiple dimensions. Their system significantly outperformed unimodal systems by providing a more complete emotional profile of users. Similarly, a recent study by Kurniawan et al. (2021) [9] integrated facial and voice analysis to develop a stress-detection system that could be deployed in real-time, demonstrating the feasibility of multimodal systems in mental health applications.

The integration of these technologies into mental health monitoring systems is also evident in the growing number of mobile applications designed for mental health care. For instance, the “Wysa” [10] app uses a combination of voice input and text-based analysis to track mental health symptoms and provide real-time feedback. The integration of AI-based techniques for emotional tracking is proving increasingly useful in the area of telemedicine and personal health monitoring.

III. METHODOLOGY

This project's methodology takes an organized approach to gathering, combining, and evaluating various data sources in order to identify mental health issues and stress. This multimodal system recognizes stress, anxiety, and depressive symptoms using voice analysis, facial expression identification, and natural language processing (NLP). A thorough explanation of each phase in the process is provided below:

A. Data Acquisition

Acquiring pertinent datasets for the system's testing and training is the initial stage in this endeavor. Datasets like FER-2013 (Facial Expression Recognition 2013) and AffectNet, which offer annotated photos of various emotional states, are utilized for facial expression recognition. These datasets aid in the system's learning of stress-related facial expressions, such furrowed brows and stiff jawlines. RAVDESS and CREMA-D datasets, which contain recordings of different emotional states exhibited through speech, such as elevated voice pitch or quivering tones that commonly accompany tension or anxiety, are used for voice analysis. In conclusion, a personalized lexicon derived from linguistic research would encompass expressions and terms that convey stress, including "overwhelmed," "anxious," or "nervous." Accurate stress detection is based on the integration of visual, aural, and language input.

B. Preprocessing

The preprocessing step ensures that the data is prepared for analysis after it has been gathered. Datasets like FER-2013 (Facial Expression Recognition 2013) and AffectNet, which offer annotated photos of various emotional states, are utilized for facial expression recognition. To achieve this, detection algorithms that find and separate the face from the backdrop, such MTCNN or Haar Cascades, are used. Subsequently, the photos undergo size normalization and grayscale conversion to ensure uniform input for the facial recognition algorithm. Preprocessing for voice data entails first eliminating background noise to guarantee distinct speech signals and then extracting characteristics such as MFCCs (Mel-frequency Cepstral Coefficients), which identify the user's emotional tonality. In order to prepare the spoken words for additional analysis in the NLP module, they are also transformed to text using a voice recognition engine such as Google voice-to-Text API.

C. Feature Extraction

Extraction of characteristics that offer significant insights into the user's emotional state comes next after preprocessing. Two different kinds of characteristics are retrieved for face expressions. Geometric features refer to the analysis of facial landmarks, including the location of the lips, eyebrows, and eyes. Stress can be indicated by certain motions, such as contracting the lips or furrowing the brow. In the meanwhile, the emotion behind the expression is classified with the use of Convolutional Neural Networks (CNNs), which automatically extract high-level information from the face pictures. Pitch, tone, and loudness are examples of prosodic features that are retrieved for voice analysis because they might indicate emotional shifts like stress or irritation. Because LSTM (Long Short-Term Memory) networks can manage the sequential nature of audio signals and identify patterns that indicate stress, they are very helpful for evaluating speech data. Using NLP techniques for the spoken words, sentiment analysis is carried out using tools like VADER or TextBlob, and models like BERT (Bidirectional Encoder Representations from Transformers) are used to assess the words' context.

D. Classification and Detection

. After characteristics are retrieved, several classification models are fed the data to identify emotions. Convolutional Neural Networks (CNNs) are used for face expression detection, and they are quite good at categorizing photos according to emotional indicators such as stress, happiness, or sorrow. For voice data, Recurrent Neural Networks (RNNs) with LSTM are used to capture the temporal relationships in the audio and classify emotions based on vocal patterns. A Support Vector Machine (SVM) classifier is used for text-based emotion detection in order to identify keywords connected to stress and the overall emotional mood. The system employs multimodal fusion to integrate the findings of the analysis of each modality—text, speech, and facial expression—after it has been completed, resulting in a more thorough and precise evaluation of the user's emotional state. The system raises a warning with greater certainty if all modalities show symptoms of stress.

D. Real time detection and feedback

The system works in real-time, constantly scanning the user's face, voice, and speech for indications of tension or anxiety in order to guarantee practical usage. A webcam's video and a microphone's audio are recorded, preprocessed, and instantly evaluated so that the system can respond right away. When the system notices indicators of stress, it alerts the user right away and provides individualized corrective actions depending on the user's current emotional state. This real-time input guarantees that the system not only recognizes mental health problems but also reacts fast to reduce the user's stress levels through treatments.

E. Stress Relief and Remedial Measures

The system offers users several corrective techniques to assist ease their symptoms when stress is recognized. One such intervention is the use of guided breathing exercises, such Box Breathing or the 4-7-8 breathing technique, which both aim to reduce anxiety by lowering heart rate and breathing rate. The system can also recommend progressive muscle relaxation (PMR), which involves the user methodically tensing and relaxing different muscle groups to reduce physical stress, in addition to breathing exercises. Additionally, mindfulness and meditation practices are provided, such as quick 5-minute mindfulness breaks or guided meditation, to assist users in removing themselves from stressful situations. Additionally, the system can recommend thought reframing influenced by cognitive behavioral therapy (CBT), which encourages users to reframe unfavorable ideas that can be elevating their stress levels. The system may recommend light exercises like stretching or walking if it determines that the user needs more exercise. These activities have been known to lower stress hormones. In more extreme situations, the system may suggest emotional self-regulation tools or professional mental health resources, making sure users have access to the help they require to handle their stress.

F. Evaluation and Testing

Using measures like accuracy, precision, recall, and F1 score, comprehensive testing is done to assess the correctness and reliability of the system. The system is customized to provide precise detection results in real-world settings by running experiments on a variety of datasets and user scenarios. In order to improve the algorithm over time and modify the suggestions and detection accuracy, user feedback will also be crucial.

G. Database and Storage

The system uses Amazon S3 in conjunction with MongoDB to handle and store user data. Real-time data, including identified stress levels, emotional states, and customized suggestions, are stored in MongoDB. For huge media assets, such audio and video data, Amazon S3 is used, guaranteeing that the system can manage enormous volumes of data while retaining speed and scalability.

IV. SYSTEM ARCHITECTURE

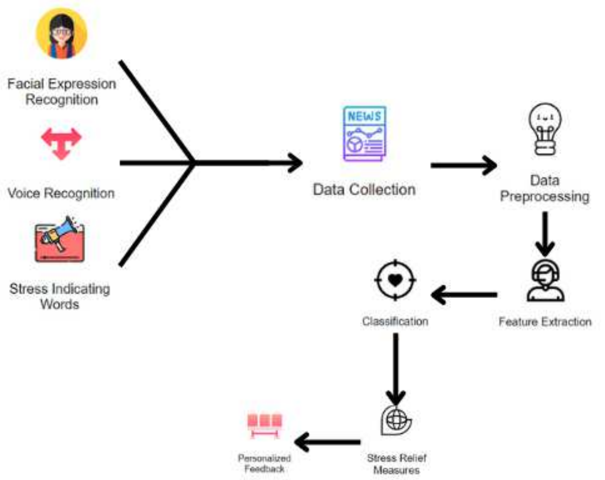

Fig 1: System Architecture of Mental Health Tracker

V. EXPECTED RESULT AND DISCUSSION

A. Expected Result

It is anticipated that our project would accurately identify early indicators of mental health conditions including stress, anxiety, and depression. The system will recognize, record, and evaluate emotional indicators using speech and facial recognition technology. Emotions such as tension, anger, fear, and sorrow may be detected with ever-greater accuracy because to the system's machine learning models, which are constantly updating. Real-time monitoring capabilities are expected to be included in the system, enabling prompt feedback and a seamless user experience. This will guarantee that users may get quick understanding of how they are feeling emotionally, which will promote early action.

Additionally, depending on the identified stress level, the tracker is anticipated to offer individualized stress-relief measures. These measures might include ideas for mindfulness activities, guided breathing exercises, or mental health resources. Personalizing feedback according to user requirements is a crucial component that makes it possible for the system to successfully handle a range of stress-related problems. Furthermore, the system may improve the depth of emotional analysis and provide a more thorough picture of the user's mental health by analyzing the user's statements that indicate stress.

In terms of usability, people with different degrees of technical proficiency may readily engage with the system because it is made to be accessible and user-friendly. Over time, it is anticipated that users would find the tool useful for controlling stress and keeping an eye on their mental health, which will boost engagement.

B. Discussion

Being able to identify mental health concerns early is one of our project's main benefits. In order to stop disorders like anxiety and depression from getting worse, early identification is essential. Unlike traditional methods of mental health evaluation that require filling out questionnaires or surveys, the system is non-intrusive since it relies on natural inputs, such as voice and facial expressions. This allows for a smooth experience. By motivating people to take proactive measures in regulating their emotions, this convenience along with the individualized nature of the feedback might greatly enhance mental health results.

Because of its architecture, the system may be scaled and adapted to function in a variety of environments, including offices, schools, and medical facilities. The emotional detection and stress-relief functions will become more accurate and effective with continued use and learning.

VI. IMPLEMENTATION

The initial implementation phase of the mental health tracking project focused on detecting stress by analyzing spoken words. This method was created in Python, utilizing the power of tools for natural language processing like NLTK and combining Flask to produce a user-friendly API endpoint. Speech transcription is the first step in the process, in which the system records voice input and turns it into text. Speech Recognition and pyaudio packages make this procedure possible, allowing for smooth real-time voice conversion and capture. Following speech transcription, the text is subjected to natural language processing (NLTK) to detect the presence of particular terms that indicate stress. This analysis is based on a thorough list of preset stress-related keywords, such as "stressed," "anxious," and "overwhelmed." After parsing the transcribed text, the algorithm checks each word against this list and flags the input if it matches, suggesting that there may be stress. The detection program counts and reports the incidence of these keywords, effectively highlighting whether stress is observed. The algorithm flags the input as suggestive of stress and offers a clear result if the user's voice contains such words.This implementation is an important milestone in the development of an approachable mental health assessment tool since it facilitates non-intrusive, real-time monitoring and feedback. It lays the groundwork for incorporating increasingly sophisticated techniques like facial expression recognition and voice tone analysis, enhancing the system's capacity to provide thorough mental health assessments.

Fig 2: Stress Indicator Detection Snap-1

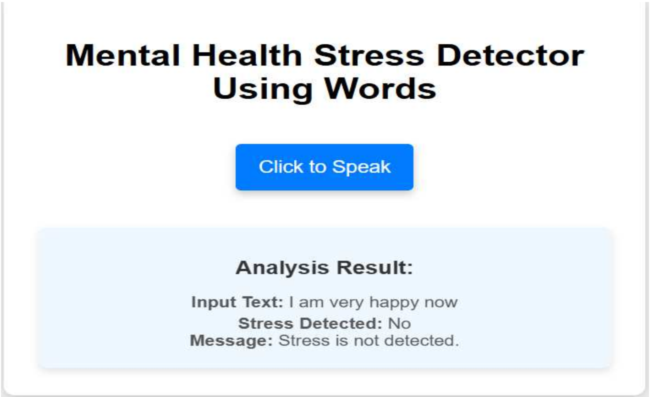

Fig 3: Stress Indicator Detection Snap-1

VII. CONCLUSION

Finally, by combining voice recognition, facial expression analysis, and the identification of phrases that convey stress, the Mental Health Tracker offers a potentially effective method of improving emotional well-being. The technology offers a user-friendly, non-intrusive means of identifying early indicators of mental health conditions including stress, anxiety, and depression by utilizing machine learning and real-time monitoring. Giving people individualized stress-relief strategies depending on their emotional state might enable them to take charge of their mental health and perhaps lessen the severity of their diseases. Although the system exhibits great promise, it is critical to recognize the obstacles pertaining to environmental accuracy, cultural sensitivity, and data privacy. Resolving these issues is essential to the system's successful adoption and scalability in a variety of contexts, such as businesses, educational institutions, and healthcare settings. This initiative is a significant step toward democratizing mental health treatment by providing people with the knowledge and skills they need to become more emotionally resilient in the hectic, fast-paced world of today.

VIII. FUTURE WORK

There are several improvements that might be made to the Mental Health Tracker. One important topic is combining data from wearables, behavior, and text to provide multimodal emotion analysis. This would enable more thorough and accurate evaluations. Integration with wearable technology, such as smartwatches, may help enhance stress detection by keeping an eye on heart rate and sleep habits. By examining emotional patterns, future iterations of the software may provide AI-powered treatment recommendations that include tailored mental health activities like cognitive behavioral therapy (CBT) or guided meditations. Improving cultural sensitivity will also ensure that the system is globally applicable by making it flexible for a variety of user groups.

IX. REFERENCES

[1] P. Ekman, "Facial Expressions of Emotion: New Findings, New Questions," *Psychological Science*, 1992.

[2] FaceReader, "Automatic Analysis of Facial Expressions," Noldus Information Technology, 2020.

[3] Liu, J., et al., "Real-Time Stress Detection using Facial Expression and Eye Movement Analysis," *Journal of Affective Computing*, 2020.

[4] Zhang, Y., et al., "A Multimodal Approach to Stress Detection using Facial Expressions and Speech Analysis," *IEEE Transactions on Affective Computing*, 2019.

[5] Schuller, B., et al., "The INTERSPEECH 2013 Computational Paralinguistics Challenge: Social Signals, Conflict, Emotion, Autism," *INTERSPEECH*, 2013.

[6] Cummins, N., et al., "Detecting Depression using Speech Analysis," *IEEE Transactions on Affective Computing*, 2018.

[7] Yang, D., et al., "Fatigue Detection in Workers using Voice Analysis," *Journal of Occupational Health*, 2020.

[8] Poria, S., et al., "Multimodal Emotion Recognition Using Facial, Textual, and Audio Features," *IEEE Transactions on Affective Computing*, 2017.

[9] Kurniawan, H., et al., "Real-Time Multimodal Stress Detection System using Facial and Vocal Data," *Sensors*, 2021.

[10] Wysa, "An AI-powered Mental Health Chatbot," Touchkin Services Pvt. Ltd., 2020.

[11].M.Iswarya ,Shivakami. A; Aravindha Mira. R; Karpagam. A “Detection of Leukemia using Machine Learning”,International Conference on Applied Artificial Intelligence and Computing 2022

[12].M.Iswarya ,”Stable And Fall Recognition In EldewaryaUsing Pigeon Hole Data Reduction And Optimization Technique And Modified KNn Classifier”,International Journal of Modern Agriculture,2021.

[13].M Iswarya, S Arun Shankar, S Abdul Hameed,”Fake Logo Detection” 2022 1st International conference on computational Science and Technology(ICCST).

[14] M.Iswarya, C Saikarthick, K Sudarshan,”Leave Management portal” 2022 1 st International Conference on Computational Science and Technology(ICCST).

[15] M. Iswarya, VS Aravinthbala.’AIB’s Gesture Language Convertor’ 2022 1st International Conference on Computational Science and Technology (ICCST).