IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A HYBRID FRAMEWORK FOR REAL-TIME STRESS DETECTION VIA BIO-SIGNAL AND AFFECTIVE ANALYSIS

1st Labeeba N
Dept. of Biomedical Engineering
KMCT College of Engineering for Women
Calicut, India

2nd Adithya P
Dept. of Biomedical Engineering
KMCT College of Engineering for Women
Calicut, India

3rd Ashna K
Dept. of Biomedical Engineering
KMCT College of Engineering for Women
Calicut, India

4th Shilpa V
Dept. of Biomedical Engineering
KMCT College of Engineering for Women
Calicut, India

5th Krishnaja P
Dept. of Biomedical Engineering
KMCT College of Engineering for Women
Calicut, India

***Abstract*—The rising problem of stress now affects mental health, work efficiency, and total human health. The process of detecting stress at an early stage works to establish both prompt treatment and the avoidance of extended health issues. The research presents a dual system which detects stress in real time through two different approaches that track body functions and assess emotional states to achieve better stress measurement outcomes. The system uses multiple methods to assess how body functions and human emotional signals interact which helps to measure a person's stress level more accurately. The research team applied machine learning and deep learning methods for their study of stress pattern detection using data that they collected. The system uses multiple models that work together to evaluate incoming data and determine whether a person experiences stress through a final stress assessment. The proposed framework uses information from multiple sources to solve the problems that single-modality systems face when detecting stress. The created method offers an effective method which can monitor people continuously while it supports mental health monitoring and workplace wellness programs and human-computer interaction needs..**

***Index Terms*—Stress detection, multimodal analysis, electrocardiogram (ECG), affective computing, deep learning, sentiment analysis.**

I. INTRODUCTION

The development of stress as a widespread problem throughout contemporary society now causes people to experience problems with both their mental and physical health. The body experiences various health issues because of persistent stress which results in anxiety disorders and cardiovascular diseases combined with decreased work efficiency and faulty decision-making abilities. The detection of stress should occur during its initial development phase because this process enables healthcare professionals to provide immediate treatment while avoiding future health issues. The process of stress identification remains difficult because it requires understanding how people react to stress through their body movements and facial expressions which differ from one individual to another. The traditional stress detection systems depend on either physiological signals or behavioral cues as their only data source. The methodologies deliver valuable information yet their effectiveness diminishes because of the intricate nature of human emotional expressions and body reactions. The use of electrocardiogram (ECG) signals allows detection of stress-related heart activity changes however these signals do not provide complete understanding of a person's emotional condition. The methods that rely exclusively on facial expressions and speech patterns and text analysis do not succeed in finding physiological stress responses. The single modality approach restricts both precision and dependability throughout the process of stress detection system assessment.

The emergence of machine learning and deep learning technologies has created intelligent systems that can study

Identify applicable funding agency here. If none, delete this.

intricate patterns through multiple data streams. The multi modal approach which merges physiological signals together with affective computing methods has demonstrated strong capabilities to enhance the accuracy of emotional state identification. Through different modalities, users achieve the capability of accessing complete information.The stress detection methods that currently exist show limitations because they depend on only one type of data which prevents them from understanding how people express their feelings through their bodily reactions. This limitation reduces the reliability of stress assessment especially in situations that require assessments to be done immediately. Our study aims to create a hybrid multimodal system that detects stress in real time through the combination of biosignal measurements and affective analysis methods. The proposed system aims to analyze multiple human cues simultaneously and utilize machine learning and deep learning models to determine whether an individual is experiencing stress. The system combines various information sources into a single framework to enhance the stress detection system's accuracy and robustness and its ability to function in real-world situations. The performance of emotion recognition systems has reached a new level of enhancement through the recent development of deep learning techniques and transformer-based architecture systems. Researchers use large dataset training methods to create models that can learn how to identify complex patterns and structures in image data and audio data and text data. Facial emotion recognition models use datasets like FER-2013 to train their systems which can identify various emotional facial expressions through facial image analysis. Wav2Vec2 speech processing framework provides a solution for extracting essential audio features needed to classify emotions from unprocessed audio recordings. DistilRoBERTa and other transformer-based natural language processing systems can deliver precise sentiment evaluation results on written materials. The current technologies enable developers to build intelligent systems which can handle different data types at the same time.

The proposed framework uses physiological data together with behavioral data to create a multimodal system which detects stress in real-time. ECG signals are acquired through electrodes using an ECG sensor module, and the signals are preprocessed to improve their quality before being analyzed. Facial expressions and speech input and textual content provide emotional cues which deep learning models use to extract emotional data from their trained datasets. The trained models operate within a unified system which handles real-time data to decide whether a person experiences stress or not.The proposed approach uses biosignal analysis together with affective computing techniques to create stress detection systems which achieve better performance through increased accuracy and improved system resilience and user-friendly design. The systems enable stress detection which can be used in multiple fields such as mental health assessment workplace wellness evaluation and human-computer interaction systems to provide early stress detection which enables timely support and better overall health outcomes.

Another important requirement for modern stress monitoring systems is the ability to perform real-time analysis. The existing research methods concentrate mainly on offline data analysis, which operates after researchers have collected physiological and behavioral data. The methods provide beneficial information yet they lack suitability for situations which need instant feedback and ongoing observation. Real-time systems help people detect stress conditions which happen in real time so they can respond at the proper time which proves essential for healthcare monitoring and workplace safety and human-computer interaction environments.

The present work proposes a hybrid multimodal framework for real-time stress detection which combines physiological biosignals with affective analysis techniques to solve these problems. The system combines information from ECG signals and facial expressions and voice inputs and textual sentiment to create a complete assessment of a person's emotional condition. The framework processes these inputs together to monitor both physiological responses and stress-related behavioral patterns.

The system uses machine learning and deep learning models to automatically detect essential patterns within the gathered data. The system processes each modality through distinct models which generate results to determine if a person is experiencing stress. The multimodal approach increases system reliability because it decreases the drawbacks which single-modality systems face.

II. RELATED WORK

Researchers now study stress detection because it serves as a major research field in both healthcare monitoring and affective computing studies. Researchers have investigated multiple methods to detect stress through human physiological signals and behavioral data and machine learning systems. Stress detection research commonly uses physiological signal analysis as its primary detection technique. The electrocardiogram (ECG) signal delivers critical data that shows how the autonomic nervous system functions. Changes in heart rate patterns can show how emotional states and stress levels develop. Researchers have used machine learning techniques to analyze ECG signals for discovering stress-related phenomena. Systems that depend only on bio-signals to deliver physio- logical data lack complete emotional expression detection for stress evaluation.

Researchers have conducted extensive research on facial emotion recognition to develop a method that detects human emotional expressions. Deep learning models use facial image datasets like FER-2013 to train for classifying emotions of anger happiness sadness fear and surprise. The models use facial feature analysis to identify emotional expressions during live analysis. Facial analysis can successfully detect emotional displays that people show but the method suffers from accuracy problems when lighting changes or head positions vary or people try to hide their feelings. Speech-based emotion recognition serves as an effective technique to identify stress levels in individuals. People express their emotions through their voice from which others can perceive their feelings by listening to their pitch and tone and voice volume. Researchers have developed advanced deep learning systems which use Wav2Vec2 technology to extract speech features from unprocessed audio data. The models achieved better results when they identified emotions through voice analysis of recorded audio.

The process of sentiment analysis for textual data has become a common method to detect emotional expressions that people show through their spoken words. DistilRoBERTa and other transformer-based models demonstrate exceptional capabilities in natural language processing which includes tasks for emotion and sentiment identification. Despite achieving success through their independent methods these systems must use multiple modes of operation since their current design limits them from properly identifying intricate stress patterns. The current study introduces a multimodal hybrid system which combines biosignal assessment with emotional state evaluation to deliver dependable and complete stress assessment results in real time.

III. PROPOSED METHODOLOGY

The proposed system introduces a hybrid multimodal framework for real-time stress detection by integrating both physiological biosignals and affective behavioral analysis. Human stress responses display complexity because people show stress through different physiological and behavioral signs. The body shows stress through physiological changes that include heart activity variations which show how the autonomic nervous system reacts to stress. People show their emotional state through facial expressions and speech tone and their choice of words. Therefore, relying on a single modality for stress detection often results in reduced accuracy and reliability.

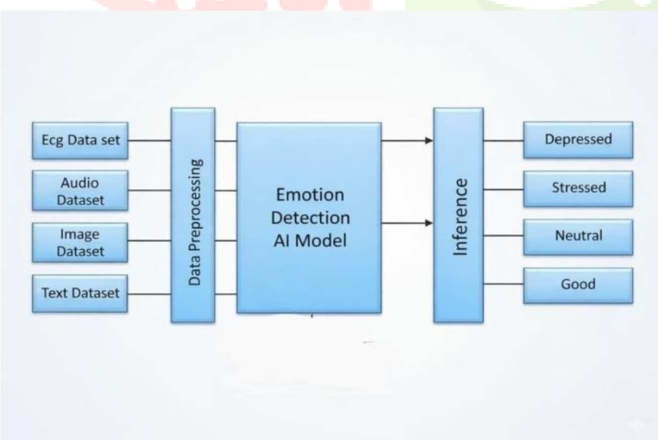

Fig. 1. Block Diagram

The proposed framework uses four different input modalities that include electrocardiogram (ECG) signals and facial expressions and voice signals and textual sentiment analysis to solve these problems. People show stress through their different stress response components which each modality detects. The ECG signal provides objective physiological measurements, whereas facial expressions, voice signals, and text analysis capture emotional and behavioral characteristics associated with stress.The system architecture consists of multiple stages including data acquisition, signal pre processing, feature extraction, deep learning-based model training, multimodal fusion, and real-time stress classification. The system processes each modality through special modules which extract essential features and determine emotional states. The system uses output from these modules to create a final binary classification which shows if the person is experiencing stress.The multi modal approach creates better stress detection because it unites physiological and behavioral indicators into one complete system.

A. Multimodal Data Acquisition

Human body ECG signals are obtained through electrode placement at designated body sites. The electrodes detect the electrical impulses that the heart produces during every heartbeat. The electrodes capture signals which have extremely low power levels that need specialized equipment to be correctly acquired and processed. Facial expression data is obtained through a camera which continuously records images of the user's face. Facial muscle movements reflect emotional states, and changes in facial expressions can provide valuable indicators of stress. A facial emotion recognition module processes these facial images to identify the emotional patterns that people show when they experience stress.

A microphone system captures voice signals. Speech signals carry emotional information which vocal characteristics such as tone pitch rhythm and intensity express. Stress often affects speech patterns, which makes voice analysis an effective tool for detecting emotions. The analysis process examines both the acoustic characteristics of speech and the development of spoken content into written text through speech-to-text technology. Natural language processing methods are used to analyze the generated textual data, which enables the system to conduct sentiment analysis and identify the emotional meanings expressed through language.The system needs simultaneous physiological and behavioral data acquisition to create a complete stress state representation of the individual.

Modality	Model / Technique
ECG Signal	AD8232 + Threshold Method
Facial Expression	HaarCascade + FER-2013
Voice Emotion	Wav2Vec2ForSpeechClassification
Text Sentiment	DistilRoBERTa

Table . 1 Models Used in the Proposed System

B. ECG Signal Acquisition and Hardware Processing

The proposed stress detection system requires physiological monitoring because stress causes changes in the autonomic nervous system which controls cardiovascular functions. ECG signals display the heart's electrical activity which occurs during its contraction and relaxation phases. Stress-related physiological changes become visible through ECG signal changes.The system uses electrodes to gather ECG signals from the AD8232 ECG sensor module in its operational process. The electrodes identify heart-generated electrical signals which they transform into analog electrical output. The signals from this system produce extremely low power output which needs to undergo amplification before any subsequent testing.The AD8232 module contains an instrumentation amplifier which functions as a specialized device to boost weak biopotential signals. The instrumentation amplifier increases the ECG signal strength without producing any distortion which would harm signal quality. The ECG signals undergo amplification to reach sufficient strength for digital systems to process them.

ECG signals experience high susceptibility to various types of noise and interference which originate from body movements and muscle contractions and electromagnetic interference and environmental sound. The AD8232 module uses filtering mechanisms to eliminate unwanted noise components which affect the signal.The integrated filtering stage helps to reduce unwanted noise frequencies while maintaining essential signal components despite the fact that amplification can create noise. The system produces a pure amplified ECG signal which accurately shows the electrical patterns of the heart.The ESP32 microcontroller transmits the filtered and conditioned signal to the computational system after the signal undergoes its initial processing. The ESP32 functions as the communication link that connects the hardware sensor module to the processing unit. The device processes analog ECG signals to create digital data which it transmits to the central system for advanced evaluation.This hardware configuration enables reliable acquisition and transmission of physiological data for stress detection.

C. ECG Signal Analysis and Stress Classification

The process begins with the acquisition and conditioning of ECG signals and then continues with the investigation of the signals to find stress-related body changes. The body experiences stress which leads to changes in the autonomic nervous system that result in alterations of both heart rate and heart rhythm. The detection of these changes can be achieved through the analysis of ECG signals. The system implements a stress detection method that uses comparators for its operation. The system establishes a threshold value which determines the normal limits of ECG signal characteristics. The system continuously checks incoming ECG signal values against the established threshold value through its comparator system.

The system interprets any ECG signal value that exceeds the threshold as an indication of potential stress. The system identifies the person as not stressed when their signal value stays underneath the established threshold.The ECG signals are used for training machine learning and deep learning models through the threshold-based method which serves as a second training method. Deep learning models can learn essential data characteristics directly from unprocessed data because they eliminate the need for manual feature identification.The trained models examine ECG patterns to discover intricate connections between physiological signals and stress conditions. The system improves stress detection accuracy through this ability which enables it to adjust for different body responses that people experience.

D. Facial Emotion Recognition Module

Facial expressions serve as the main method which people use to show their emotional states. Psychological studies have shown that people show their emotional states through facial muscle movements which display anger and fear and sadness and happiness. The proposed multimodal stress detection system requires facial emotion recognition as its fundamental part.The facial emotion recognition system operates through a deep learning model which has been trained on the FER-2013 dataset that includes about 7000 facial images with different emotion categories.The dataset includes seven emotion classes:

- Angry
- Disgust
- Fear
- Happy
- Neutral
- Sad
- Surprise

One challenge associated with this dataset is class imbalance, particularly in the disgust category where the number of samples is significantly lower compared to other classes. The solution to this problem uses data augmentation techniques which create new dataset variations through their respective methods. Data augmentation creates new image content by applying various transformations to existing images which include rotation and flipping and scaling and cropping and brightness adjustments.The system requires facial detection to identify face locations within the captured image before it can proceed with emotion classification.This is accomplished using the Haar Cascade face detection algorithm, which is a machine learning-based method designed to detect objects within images. The Haar Cascade algorithm scans the image and identifies facial regions based on predefined features.Once the face is detected, the facial region is extracted and passed to the trained deep learning model, which predicts the emotional state associated with the detected facial expression.The predicted emotion is then forwarded to the multimodal fusion module for final stress analysis.

E. Voice Emotion Recognition Module

The speech signals contain multiple emotional signals which psychologists can use to assess a person's mental condition. People under emotional stress show changes in their vocal patterns which include alterations in pitch and tone and speech rate and vocal intensity. Voice signal analysis helps researchers understand human emotional states.The system performs speech emotion recognition through a deep learning framework which uses the Wav2Vec2 architecture. The system uses the Wav2Vec2FeatureExtractor, which is a preprocessing component from the Hugging Face Transformers library. The feature extractor transforms raw audio data into numerical formats which deep learning systems can execute.

The system uses Wav2Vec2ForSpeechClassification as its speech emotion classification architecture while Light Eternal serves as its main base model. The model classifies speech signals into five emotional categories which include anger disgust fear happiness and sadness. Audio signals are processed through the Librosa library, which enables users to load audio files and extract audio features and prepare audio data for their analytical work. The recorded speech signals are segmented into smaller chunks before classification to improve model accuracy and ensure efficient processing.The model evaluates each audio segment to determine which emotional category it belongs to. The multimodal stress detection module uses these predictions as its input.

F. Text Sentiment Analysis Module

In addition to analyzing acoustic properties of speech, the spoken content is converted into text and analyzed using natural language processing techniques. Textual sentiment analysis helps identify emotional intent conveyed through language.The sentiment analysis module uses the DistilRoBERTa-base transformer model which provides a lightweight and efficient implementation of the RoBERTa architecture that developers created for sequence classification tasks.The architecture used is RobertaForSequenceClassification, which processes textual input and classifies it into multiple emotional categories. The model predicts labels such as anger, disgust, fear, joy, neutral, sadness, and surprise.The text sentiment analysis module enables the system to detect emotional signals which people express through their language because these signals do not always show through their facial expressions or voice characteristics

G. Multimodal Model Integration and Real-Time Inference

In addition to analyzing acoustic properties of speech, the spoken content is converted into text and analyzed using natural language processing techniques. Textual sentiment analysis helps identify emotional intent conveyed through language.The sentiment analysis system uses DistilRo BERTa base transformer model which functions as a compact and effective implementation of the RoBERTa design used for processing sequential data. The system employs RobertaForSequenceClassification architecture to analyze written content and determine its emotional classification across various categories. The model predicts labels such as anger, disgust, fear, joy, neutral, sadness, and surprise.The text sentiment analysis module allows the system to capture emotional signals which people express through language because these signals do not always show through their facial expressions or voice characteristics.

H. User InterfacOutput Visualization

The final output of the proposed system is displayed through a user interface designed for real-time monitoring. The interface receives the prediction generated by the integrated model and presents the result in a clear and simple format. The interface assesses an individual's stress level based on the classification result which the multimodal model provides. The interface presents the main output together with additional information which includes facial analysis results of detected emotions and voice emotion recognition outcomes and sentiment analysis findings.The interface enables continuous updates because it processes fresh data which allows users to monitor stress levels in real time. The interface uses visual alerts and warning displays to inform users when the system identifies a stress condition. The system provides value to applications that include health monitoring and workplace stress management and emotional wellbeing assessment.

I. Model Performance Evaluation

The performance of the individual models used in the proposed multimodal stress detection system was evaluated based on classification accuracy.The facial emotion recognition model was trained using the FER-2013 dataset with face detection performed using the Haar Cascade algorithm. This model achieved an approximate accuracy of 65percentage in detecting facial emotions.The voice emotion recognition model was developed using the Wav2Vec2ForSpeechClassification architecture with the LightEternal model. The model processed audio signals through the Librosa library which resulted in an accuracy performance of 75percentage.The DistilRoBERTa-base model with RobertaForSequenceClassification architecture was used for text sentiment analysis.This model classified the emotional content of text with an approximate accuracy of 66percentage.Although individual models provide moderate accuracy, combining these modalities improves the overall reliability of the stress detection system.

IV. RESULT AND DISCUSSION

The performance of the proposed hybrid multimodal stress detection system was evaluated by analyzing the outputs produced by each individual module, including facial emotion recognition, voice emotion recognition, text sentiment analysis, and ECG-based stress detection. Each module was trained using appropriate datasets and deep learning architectures. The results obtained from these models demonstrate how different modalities contribute to identifying emotional and physiological indicators of stress. The following subsections describe the results obtained from each modality and discuss their effectiveness.

A. Facial Emotion Recognition

The facial emotion recognition module was developed using the FER-2013 dataset which contains thousands of grayscale facial images that are categorized into seven emotional classes which are angry disgust fear happy neutral sad and surprise. The system starts face detection through the Haar Cascade algorithm which works by detecting facial areas from live webcam and camera module images.The trained deep learning model receives the cropped facial image after the system detects the face region to determine the emotional state of the individual. The researchers used data augmentation techniques that included image rotation and flipping and scaling during training to create a diverse dataset which would enhance the model's ability to generalize.

The experimental results showed that the facial emotion recognition model achieved an approximate accuracy of 65 percentage. The model performance was restricted by three factors which included different lighting conditions and face partial occlusion and human facial expression variations. The FER-2013 dataset exhibits a class imbalance issue specifically in the disgust category which decreases the model's capacity to accurately identify particular emotions. The facial emotion recognition module successfully detects all visible emotional cues which include facial muscle movements and eye expressions and mouth curvature. The features deliver essential behavioral data which assists in detecting emotional patterns associated with stress.

B. Voice Emotion Recognition

The voice emotion recognition module analyzes speech signals to detect emotional patterns which people exhibit when they experience stress. Human speech carries important emotional information through variations in pitch and tone and intensity and rhythm. The vocal features of a person show major differences because of his or her emotional states which include stress and anger and sadness.The system requires users to record their speech through a microphone which Librosa audio processing library uses to process the audio signals. The system divides the recorded audio signal into smaller segments which help with feature extraction and classification tasks.

The Wav2Vec2FeatureExtractor processes extracted audio signals by transforming their raw waveform signals into usable formats for deep learning models.The classification model used for voice emotion recognition is LightEternal with the Wav2Vec2ForSpeechClassification architecture.The experimental test results demonstrated that the voice emotion recognition module achieved 75 percentage accuracy which made it the most precise method among the behavioral inputs used in this system. The model demonstrates better performance because vocal features serve as effective indicators for recognizing stress-related emotional states.The voice-based model experiences performance changes because of three factors which include background noise microphone quality and speech clarity. Speech analysis continues to be an effective method for detecting emotional distress across various real-world situations.

C. Text Sentiment Analysis

The text sentiment analysis module evaluates the emotional meaning of spoken or written language.The system starts with voice input which gets transformed into text through speech-to-text technology. The text undergoes analysis through natural language processing methods which determine the emotional sentiment that the user has expressed.The system employs the DistilRoBERTa-base transformer model together with the RobertaForSequenceClassification architecture to perform text sentiment classification. DistilRoBERTa provides a lightweight alternative to RoBERTa which enables users to understand language without needing extensive processing power.The model was trained to classify text into seven emotional categories: anger, disgust, fear, joy, neutral, sadness, and surprise. These categories help in understanding the emotional tone of the spoken or written content.

The text sentiment analysis module reached an accuracy rate of about 66 percentage during its evaluation. The model performance depends on how clear and contextualized the input text is presented to it. Short or ambiguous statements may not provide sufficient information for accurate emotion detection. The use of sarcasm and indirect expressions results in decreased accuracy for classification tasks. The text sentiment analysis module delivers essential linguistic information about a person's emotional state while it enhances the system through its combination with existing assessment methods.

D. ECG-Based Stress Detection

The ECG-based stress detection module shows the user's stress level through its physiological measurements. The ECG system collects heart activity electrical signals through electrodes that connect to the AD8232 ECG sensor module.The raw ECG signals obtained from the electrodes contain noise which results from three sources: motion artifacts and power line interference and muscle activity. The AD8232 module uses an instrumentation amplifier for its solution because the amplifier increases weak ECG signals while it removes both signal and noise interference.

The ESP32 microcontroller receives the amplified signal which it sends to the system for analysis after processing the data.In this system, a threshold-based comparator method is used to determine stress levels from ECG signals. A predefined threshold value is established based on physiological signal patterns. If the measured ECG parameter exceeds the threshold value, the system identifies the user as stressed. Otherwise, the user is classified as not stressed.Although the ECG-based approach is relatively simple compared to deep learning methods, it provides objective physiological evidence of stress that complements behavioral data obtained from facial, voice, and text analysis.

V. CONCLUSION AND FUTURE WORK

This research developed a hybrid multimodal framework which detects stress in real-time through the combination of physiological signal monitoring and affective computing methods. The system integrates four distinct input modalities which include ECG signals, facial expressions, voice signals and textual sentiment analysis. The system analyzes multiple human cues at once to improve stress detection accuracy beyond the capabilities of traditional methods which depend on single data sources. The system used AD8232 sensor

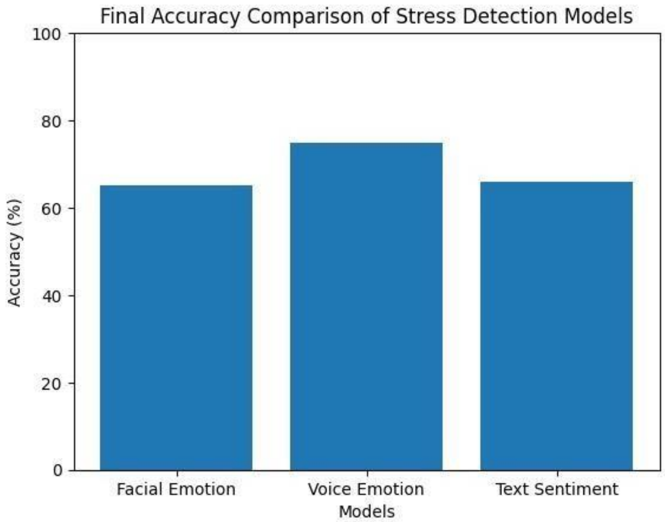

Fig. 2. Accuracy comparison of stress detection models.

module electrodes to record ECG signals which the module amplified and filtered to send to the ESP32 microcontroller processing unit. The physiological signals provide objective data that shows how heart activity changes when a person experiences stress. The researchers used machine learning and deep learning models to investigate behavioral cues of the subjects. The model used to detect facial emotions combined the FER-2013 dataset with the Haar Cascade algorithm to detect faces. The system used Wav2Vec2ForSpeechClassification for voice-based emotion recognition and LightEternal model for voice-based emotion recognition while DistilRoBERTa model with RobertaForSequenceClassification architecture performed textual sentiment analysis. The system received output from all modules which it used to create a unified system that could determine whether a person was experiencing stress.

The experimental findings show that every stress detection method brings essential information to the detection process. The facial emotion recognition module achieved an accuracy of approximately 65 percentage, the voice emotion recogni- tion model achieved around 75percentage accuracy, and the text sentiment analysis model achieved about 66 percentage accuracy. The system becomes more dependable and stronger through multiple modality combinations which overcome the performance challenges faced by different models in various environments and with different datasets. The multimodal approach allows the system to analyze both physiological and behavioral indicators simultaneously, providing a more comprehensive understanding of a person's emotional state. The integration of these two systems helps to decrease the chances of errors which typically arise from using only one system. The proposed system shows how biosignal analysis can be combined with affective computing methods to develop intelligent systems that monitor emotional well-being through- out the day. The systems have applications for healthcare monitoring and workplace stress management and mental health assessment and assistive technology development. The system continuously tracks physiological signals and behav- ioral patterns to provide users with early stress alerts which enable them to take corrective actions before their condition worsens.

Future work may benefit from multiple enhancements which researchers should explore. The facial emotion recognition model requires improved accuracy through attendance of larger balanced training datasets which include advanced deep learning models based on convolutional neural networks with attention mechanisms. The ECG analysis component needs upgrades through the addition of more physiological features which include heart rate variability (HRV) parameters because these metrics deliver better understanding of autonomic nervous system functioning. The researchers need to investigate better ways to combine multimodal data through their current work which needs advanced methods like ensemble learning and multimodal deep neural networks to combine different data types. The system can achieve continuous stress monitoring through its integration with wearable devices and mobile health applications which will track users' stress levels throughout their daily routines. The research should investigate cloud-based processing methods which will enhance user interfaces while delivering real-time feedback that gives users personalized stress management recommendations.The multimodal stress detection system shows that using physiological signals together with affective computing techniques enables systems to better understand human emotional states. The system needs further enhancements and testing in real-life situations before it can become an effective tool for detecting stress and monitoring mental health and preventing health problems.

REFERENCES

[1] S. Gilboa, A. Shirom, Y. Fried, and C. Cooper, "A meta-analysis of work demand stressors and job performance: Examining main and moderating effects,"

[2] P. Melillo, C. Formisano, U. Bracale, and L. Pecchia, "Classification tree for reallife stress detection using linear heart rate variability analysis. Case study: Students under stress due to university examination," in Proc. World Congr. Med.

[3] M. Maes, "The cytokine hypothesis of depression: Inflammation, ox- idative nitrosative stress (IONS) and leaky gut as new targets for adjunctive treatments in depression," Neuro Endocrinology Lett., vol. 29, pp. 287–291, Jun. 2008.

[4] T. Chandola, A. Heraclides, and M. Kumari, "Psychophysiological biomarkers of workplace stressors," Neurosci. Biobehavioral Rev., vol. 35, no. 1, pp. 51–57, Sept2010.

[5] M. F. Dallman et al., "Feast and famine: Critical role of glucocorticoids with insulin in daily energy flow," Frontiers Neuroendocrinology, vol. 14, no. 4, pp.303–347, Oct. 1993.

[6] R. C. Kessler, "The effects of stressful life events on depression," Annu. Rev.Psychol., vol. 48, no. 1, pp. 191–214, Feb. 1997.

[7] s. Esler, "Mental stress and human cardiovascular disease," Neurosci.
Biobehavioral Rev., vol. 74, pp. 269–276, Mar. 2017.

[8] S. Pourmohammadi and A. Maleki, "Stress detection using ECG and EMG signals:A comprehensive study," Comput. Methods Programs Biomed., vol. 193, Sep. 2020,Art. no. 105482.

[9] R. Dantzer, "Depression and inflammation: An intricate relationship,"
Biol. Psychiatry, vol. 71, no. 1, pp. 4–5, Jan. 2012.

[10]]R. Dantzer, J. C. O'Connor, M. A. Lawson, and K. W. Kelley, "Inflammation associated depression: From serotonin to kynurenine," Psychoneuro endocrinology,vol. 36, no. 3, pp. 426–436,Apr.2011