International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Study on Machine Learning-Based Stress Recognition System

K. Sangeetha[1], A. Sureshkumar[2]

Submitted: 06/10/2024 **Revised:** 18/11/2024 **Accepted:** 06/12/2024

Abstract: Stress emerges as the body's reaction to shifts in the surroundings exert influence, manifesting through a multitude of cognitive, physiological, or affective reactions. Prolonged acute stress may disrupt both physiological and psychological well-being equilibrium, resulting in decreased work efficacy and a heightened risk of chronic ailments such as hypertension and anxiety disorders. As psychological stress increasingly becomes a global issue, impacting people of all ages, there is an urgent demand for effective monitoring systems. A dependable and economical an acute stress detection system could allow individuals to track and regulate their stress levels, thus alleviating long-term negative outcomes. This article examines and discusses literature centered on machine learning-driven strategies for stress detection, highlighting their potential for real-time oversight. Furthermore, we delve into existing solutions that incorporate edge computing technologies, improving the practicality and efficacy of stress monitoring in real-world scenarios. By amalgamating current research, this review aspires to underscore the progress in machine learning methodologies for stress detection and the significance of edge computing in delivering timely and actionable insights for stress management.

Keywords: *Stress monitoring, machine learning, edge computing, real-time detection, psychological health*

INTRODUCTION

Stress is an inherent part of human existence, triggered by various external and internal factors. Defined as the body's physiological and psychological response to changes in the environment, stress can manifest in numerous ways, impacting an individual's overall well-being. While short-term stress responses are often adaptive, prolonged exposure to stress can lead to severe health consequences. Acute stress, characterized by short bursts of high stress, poses significant risks if not monitored and managed effectively.

The prevalence of stress-related disorders is rising globally, necessitating the development of efficient monitoring systems to facilitate early intervention and management. Machine learning (ML) has emerged as an invaluable resource across various domains, particularly within the healthcare sector, where it reveals innovative techniques for the identification and monitoring of stress. By analyzing extensive datasets, machine learning algorithms are capable of identifying patterns and correlations that may escape the attention of human analysts. This capability is particularly advantageous in the realm of stress assessment, given the significant variability in individual responses to stressors. The application of machine learning methodologies facilitates the development of customized stress management systems that adapt to the unique stress profiles of individuals, thereby enhancing the overall effectiveness of stress alleviation strategies. Recent advancements in wearable technology and sensors have further propelled the integration of machine learning into real-time stress monitoring. These devices can continuously collect physiological metrics including heart rate variability, skin conductance, and body temperature, which serve as crucial markers of stress levels. When combined with machine learning algorithms, this data can be analyzed instantaneously, providing users with immediate feedback regarding their stress states. This prompt information empowers individuals to undertake proactive measures to mitigate their stress before it escalates into more serious health complications.

Stress is characterized as the physiological response of the body to unfavorable environmental stimuli that impede an individual's standard coping strategies [1]. Although positive stress (eustress)

PG Scholar[1], Assistant Professor[2]
Department of Computer Science and Engineering, Excel Engineering College,
Namakkal, Tamil Nadu 637303
Correspondence mail id: dguidephd@gmail.com
suresh308it@gmail.com

serves to enhance concentration and assists individuals in overcoming challenges, adverse stress (distress) ignites the engagement of the hypothalamic-pituitary-adrenal (HPA) axis. Prolonged activation of the HPA axis could lead to both somatic and psychological complications [2]. Moreover, psychological stress can disrupt bodily functions and diminish work performance in daily activities, potentially resulting in adverse effects on the economy [3]. Monitoring levels of negative stress can yield essential insights for recognizing stressors and executing interventions to avert future disturbances.

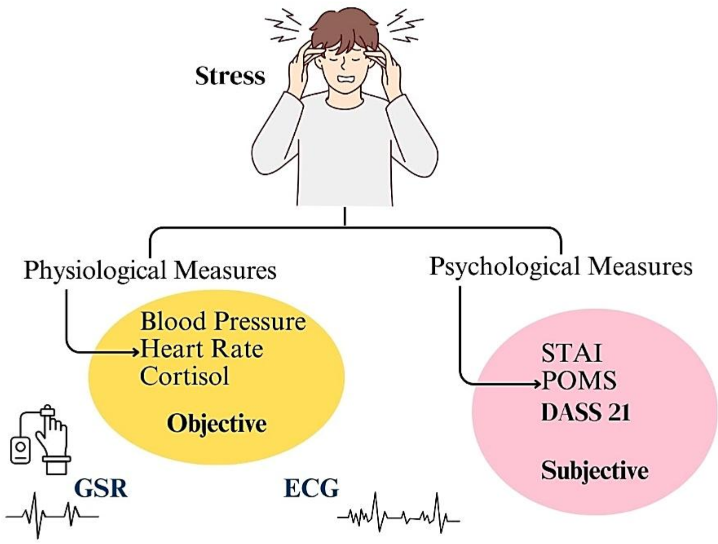

Figure 1. Examples of Objective and Subjective Stress Assessment Methods.

Stress can be divided into two unique categories: (i) physiological or "objective" stress and (ii) psychological or "subjective" stress, often called perceived stress. Objective stress reveals itself through changes in physical indicators like increased blood pressure, heart rate, and cortisol levels. In contrast, subjective stress depends on how an individual perceives the stressfulness of a specific situation. Commonly used methods for evaluating perceived stress include questionnaires such as the DASS 21 (Depression, Anxiety, and Stress Scale), STAI (State-Trait Anxiety Inventory), and POMS (Profile of Mood States) (Figure 1).

Noteworthy physiological markers of stress include

(i) cortisol (levels of the stress hormone)

(ii) Signals from GSR (Galvanic Skin Response), ECG (Electrocardiogram), and EEG (Electroencephalogram).

In reference [4], various physiological indicators of stress and the associated technologies utilized for their measurement were examined. Furthermore, [5] provided a review of diverse sensors and commercial devices designed for the assessment of stress. This article investigates machine learning methodologies for the detection of stress and evaluates the literature concerning real-time models of stress monitoring.GSR is recognized as the most widely utilized physiological measure of stress. It encompasses both physiological and psychological arousal, wherein the activation of the autonomic nervous system (ANS) amplifies sweat gland activity, thereby enhancing skin conductance. Figure 2 illustrates the correlation between GSR and the transition of the ANS from a state of stress to relaxation [6]. Nevertheless, the sole utilization of GSR for stress identification can be intricate due to issues related to signal quality and variability in responses. The reliance exclusively on GSR has proven insufficient in differentiating between varying levels of stress. For instance, [7], a combination of Electrodermal Activity (EDA) and Photoplethysmograph (PPG) signals achieved greater accuracy compared to the use of EDA in isolation [8]. Approaches employing multiple sensors typically surpass those reliant on a single sensor, indicating that stress-monitoring devices should incorporate multiple sensors for enhanced accuracy in detection. Wearable sensor systems are optimal for the real-time monitoring of stress, providing comfort, convenience, and unobtrusive observation.

Innovative prototypes and methodologies for the evaluation of stress through wearable technology have been extensively documented [9], a stress-monitoring patch capable of measuring skin temperature, skin conductance, and pulse wave signals was introduced. [10-12] describe the development of a glove equipped with EDA and pulse wave sensors. Researchers at the MIT Media Lab devised a system for the monitoring of physiological signals aimed at enhancing communication [13]. Additionally, a sensor designed to measure heart rate, skin conductance, and skin temperature was proposed [14]. [15] Details the creation of an advanced body sensor network tailored for ambulatory stress monitoring.

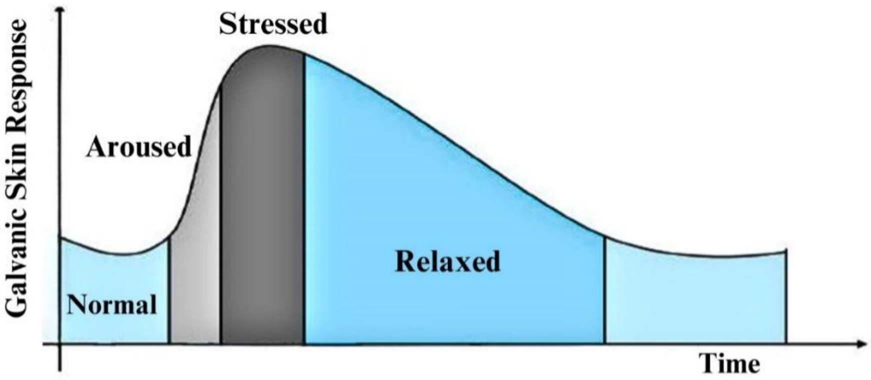

Figure 2. Fluctuations in Galvanic Skin Response in relation to psychological stress [6].

Notwithstanding, there exists a plethora of commercially accessible devices and configurations intended for the acquisition and recording of physiological signals.

METHOD

Data Acquisition

Physiological data sourced from the PHYSIONET repository (http://www.physionet.org/), crafted by the innovative minds of Jennifer Healey and Rosalind Picard, were employed for this investigation. This repository comprises information from healthy participants who navigated a route through Boston, encompassing urban streets (high tension), motorways (moderate tension), and intervals of relaxation (low tension). The compilation features 17 drivers, with seven individuals (06, 07, 08, 10, 11, 12, and 15) chosen for possessing comprehensive data.

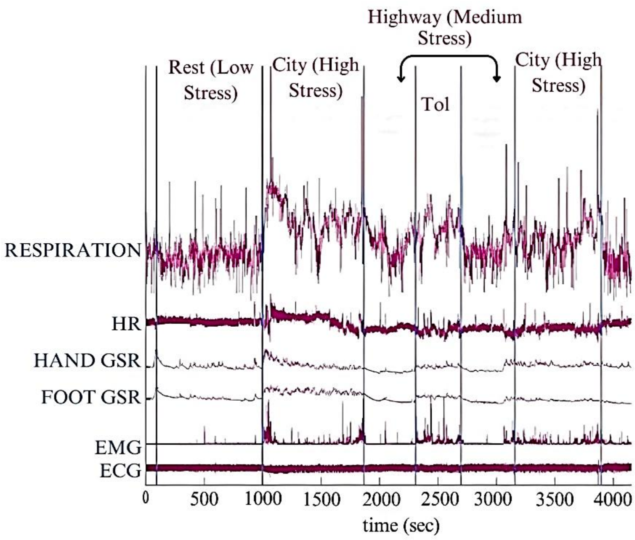

Figure 3. Different signals for 'drive06'

Five physiological indicators for every driver were examined: Foot Galvanic Skin Response (FGSR), Hand Galvanic Skin Response (HGSR), Electromyography (EMG), Heart Rate (HR) derived from Electrocardiogram (ECG) data, and Respiration (RESP). Figure 1 showcases the signals recorded for 'drive06' in the study by Healey and Picard. Preprocessing: The dataset is segmented into distinct portions that align with different stress levels for analysis purposes. Time frames of 100, 200, and 300 seconds are implemented to signify three unique stress categories: low stress (relaxation), moderate stress, and high stress. Each physiological signal is then partitioned into nine sections based on 100-second intervals. The first three overlapping sections represent the initial resting phase, identified as low stress (Figure. 2). The next three overlapping sections relate to the first urban driving phase, linked to high stress (Figure 3), while the last three overlapping segments are associated with the first highway phase, marked as moderate stress.

Feature Extraction

In each segment, we unveil 78 distinct characteristics. Every characteristic is selected from the plethora of vital and frequently utilized aspects pertaining to physiological signals, as referenced in sources [2-13]. An overview of these characteristics can be found in table 1.

Table 1 Symbolic Features and Their Descriptions

Feature Description	EMG	HR	Foot GSR	Hand GSR	RESP
Mean Normalization	EMG21	HR21	FGSR21	HGSR21	RESP21
Root Mean Square (RMS)	EMG22	HR22	FGSR22	HGSR22	RESP22
Average Power 0.01 - 0.1 Hz	EMG23	HR23	FGSR23	HGSR23	RESP23
Average Power 0.1 - 0.2 Hz	EMG24	HR24	FGSR24	HGSR24	RESP24
Average Power 0.2 - 0.3 Hz	EMG25	HR25	FGSR25	HGSR25	RESP25
Average Power 0.3 - 0.4 Hz	EMG26	HR26	FGSR26	HGSR26	RESP26
Average Power F1 - F2 Hz	EMG27	HR27	FGSR27	HGSR27	----
Average Power F3 - F4 Hz	EMG28	HR28	FGSR28	HGSR28	----
Ratio Low Band / High Band	EMG29	HR29	FGSR29	HGSR29	RESP27
Difference Between Adjacent Elements (Means)	EMG30	HR30	FGSR30	HGSR30	RESP28
Difference Between Adjacent Elements (2nd times)	EMG31	HR31	FGSR31	HGSR31	RESP29
Interquartile Range (IQR)	EMG32	HR32	FGSR32	HGSR32	RESP30
Sum of Rise Time (10% to 90% of Reference Levels)	EMG33	HR33	FGSR33	HGSR33	RESP31
Peak 2 Peak	EMG34	HR34	FGSR34	HGSR34	RESP32
Sum of Local Peak	EMG35	HR35	FGSR35	HGSR35	RESP33
Number of Local Peaks	EMG36	HR36	FGSR36	HGSR36	RESP34

In Table 1, certain frequencies remain unspecified, which we have subsequently presented in Table 2.

Table 2. Undefined Frequency in Table 1

Signal Frequency (Hz)	Frequency 1 (F1)	Frequency 2 (F2)	Frequency 3 (F3)	Frequency 4 (F4)	Low Band Frequency	High Band Frequency
Electromyography (EMG)	0.35	0.65	0.05	5.5	0.05 to 0.2	0.05 to 6.0
Heart Rate (HR)	0.03	0.25	0.22	0.55	0.03 to 0.22	0.22 to 0.6
Foot Galvanic Skin Response (Foot GSR)	0.04	0.55	0.55	1.8	0.04 to 0.55	0.55 to 1.8
Hand Galvanic Skin Response (Hand GSR)	0.04	0.55	0.55	1.8	0.04 to 0.55	0.55 to 1.8
Respiration	0.015	0.12	0.18	0.35	0.015 to 0.12	0.35 to 0.45

Feature Selection

The feature vectors encompass seventy-eight distinct attributes for each segment of all signals, resulting in extended training durations and complex computations. Consequently, the optimization of feature selection markedly improves the efficiency of classification methodologies. A feature selection algorithm generally integrates a search technique to propose new subsets of features with an evaluation criterion to assess these subsets. The most straightforward algorithm evaluates every conceivable feature subset to determine the one that minimizes the error rate. Within the Weka software environment, all features are prioritized utilizing CfsSubsetEval and InfoGain AttributeEval.Conversely, an alternative algorithm leverages machine learning techniques to discern the most pertinent features. In Weka, the optimal features for classification are established through ClassifierSubsetEval in conjunction with the SVM classifier. This investigation employs both algorithms to ascertain the most significant features. Initially, by amalgamating both methodologies, all features are ranked, with the highest selections merged with those identified by the secondary algorithm, thereby facilitating classification across multiple states. In contrast, the latter approach exclusively depends on features selected by the second algorithm, thereby constraining classification to a singular state.Subsequent to identifying the effective features, SVM and KNN classifiers with cross-validation are utilized for classification purposes. Signal processing and feature extraction are executed using MATLAB 2012a, followed by the implementation of SVM and KNN for classification in WEKA 3.6.

RESULTS

In the case of three unique conditions, the durations of 100 seconds, 200 seconds, and 300 seconds are examined. The outcomes pertaining to Support Vector Machine (SVM) and K-Nearest Neighbors (KNN), Table 3 relating to different sensor counts and varied feature collections, are detailed in tables 3, 4, 5, 6, 7, and 8.

Table 3 Assessment characteristics for intervals of 100 seconds utilizing Support Vector Machine (SVM) methodologies.

Features used for 120 seconds state	Number of Features	SVM Classification Accuracy	Number of Sensors

ALL	75	89.50%	6
RESP 1,2,6,7,8,9,10 – HR 2 – EMG 7,10,12 – HGSR 9,11,14,16 – FGSR 5,7,9,12,14	18	97.85%	5
RESP 3,6,8,9 – HR 2 – EMG 15 – HGSR 9 – FGSR 5	9	95.75%	5
RESP 3,6,8,9 – HR 2 – HGSR 9	7	92.40%	3
RESP 3,6,8,9 – HGSR 9	6	86.20%	2
RESP 1,3,6,8,12	6	83.50%	1
ClassifierSubsetEval feature selection: EMG 4,12,15 – HGSR 6,14 – FGSR 8 – RESP 5,6,9,12	12	81.30%	4

In the realm of the 100-second interval state, the pinnacle of accuracy is attained through the deployment of the SVM classifier, utilizing an extensive assortment of five sensors and twenty features. Moreover, Table 4 superior precision is realized through a meticulously selected feature set employing the SVM classifier, Table 5 which incorporates all five sensors and eight features. In instances where there is a decrease in the number of sensors and features, the KNN classifier demonstrates improved accuracy, depending solely on one sensor and three features.

Table 4 Examination characteristics over intervals of 200 seconds employing Support Vector Machine (SVM) methodology.

Features used for 150 seconds state	**Number of Features**	**KNN Classification (%)**	**Number of Sensors**
ALL	80	88.55%	6
RESP 1,4,9,12 - HR 2 - EMG 7,12,16 - HGSR 9,12 - FGSR 3,12,17	19	94.85%	6
RESP 1,5,11,13 - HR 2 - EMG 10,15 - HGSR 12 - FGSR 3	9	96.43%	5
RESP 1,12,14 - HR 2 - EMG 7,10 - HGSR 9	7	94.35%	5
RESP 1,12,13 - HGSR 9	5	89.80%	4
RESP 2,6	4	93.20%	3
ClassifierSubsetEval feature selection: EMG 5,13,17 - HGSR 10,12 - FGSR 3,7,9,12	11	91.25%	5

In the context of a 200-second interval state, the utmost precision is attained through the utilization of an SVM classifier, incorporating all five sensors and sixteen features. Conversely, when considering a limited number of features, Table 6 the highest accuracy is accomplished with a KNN classifier, utilizing four sensors and seven features. Furthermore, in scenarios involving a reduced count of sensors and features, the optimal accuracy is realized via the KNN classifier, employing a singular sensor and three features.

Table 5 Examination characteristics for intervals of 200 seconds utilizing KNN.

Features used for 150 seconds state	**Number of Features**	**KNN Classification (%)**	**Number of Sensors**
ALL	80	88.55%	6
RESP 1,4,9,12 - HR 2 - EMG 7,12,16 - HGSR 9,12 - FGSR 3,12,17	19	94.85%	6
RESP 1,5,11,13 - HR 2 - EMG 10,15 - HGSR 12 - FGSR 3	9	96.43%	5
RESP 1,12,14 - HR 2 - EMG 7,10 - HGSR 9	7	94.35%	5
RESP 1,12,13 - HGSR 9	5	89.80%	4
RESP 2,6	4	93.20%	3
ClassifierSubsetEval feature selection: EMG 5,13,17 - HGSR 10,12 - FGSR 3,7,9,12	11	91.25%	5

Table 6 Examination attributes for intervals of 300 seconds employing Support Vector Machines (SVM).

Features used for 400 seconds state	**Number of Features**	**SVM Classification (%)**	**Number of Sensors**
ALL	80	81.25%	6
RESP 1,4,7,10 - HR 2 - EMG 5 - HGSR 4 - FGSR 8,9	10	98.75%	5
RESP 2,6,8 - HGSR 7 - FGSR 4	5	86.45%	3
RESP 1,5,9	4	82.20%	2
ClassifierSubsetEval feature selection: EMG 2 - HGSR 5 - RESP 8,10	6	94.65%	4

Table 7 Examination of characteristics over intervals of 300 seconds utilizing KNN methodology.

Features used for 300 seconds state	**Number of Features**	**KNN Classification**	**Number of Sensors**
ALL	75	80.12%	6
RESP 5, 10, 12 – HR 3 – EMG 4 – FGSR 4 – HGSR 4	8	92.56%	5
RESP 2, 7, 13 – EMG 1 – HGSR 2	6	97.85%	4

RESP 1, 3, 11 – FGSR 1 – HGSR 2	5	88.63%	3
RESP 6, 7, 8	4	84.71%	2
ClassifierSubsetEval feature selection: EMG 4 – HGSR 2 – RESP 6	3	93.67%	3

In the realm of a 300-second temporal state, the pinnacle of precision is achieved through the deployment of the KNN classifier, which utilizes three sensors and five distinct features. As depicted in Tables 7 through 7, enhanced accuracy is noted with a lesser count of sensors and features during prolonged time spans, and it is significant to point out that the respiration sensor stands out as the most vital instrument for identifying stress. Ultimately, the results of this investigation, in conjunction with three other studies, are compiled in Table 8. As shown, the conclusions of this research reveal superior accuracy while leveraging a reduced array of features.

Table 8 Compare the results

Time int. (s)	Acc. (%)	Classifier	Sensor numbers	Features numbers	Ref.
250	96	k-NN	4	18	[2]
250	82.75	Decision Tree	3	6	[5, 7]
200	95.5	Random Forest	5	12	[6]
150	97.12	Naive Bayes	4	14	[8, 9]
120	99.25	SVM	5	19	current paper

CONCLUSION

In conclusion, the pressing issue of stress and its associated health risks necessitates the development of effective monitoring solutions. Machine learning has proven to be a valuable tool in identifying and analyzing stress patterns, offering promising avenues for real-time detection and management. By utilizing various algorithms, from supervised learning methods to more advance in the realm of neural networks, scholars have achieved considerable advancements in improving the precision and dependability of stress detection systems. These systems can provide valuable insights, enabling individuals to understand their stress responses better and take necessary actions to mitigate its impact. The integration of edge computing technologies further enhances the practicality of stress monitoring solutions. By processing data locally, edge computing reduces latency and ensures that users receive timely feedback on their stress levels. This immediacy is crucial, as it allows for real-time interventions that can prevent the escalation of stress into chronic conditions. Additionally, edge computing offers improved data privacy, addressing concerns about sensitive health information being transmitted to cloud servers. Together, these advancements represent a significant leap forward in creating user-centric stress management solutions.

It has been established that levels of stress can be discerned through biological indicators, employing a diverse array of biological sensors, unique characteristics, and varying timeframes. The most efficacious attributes are chosen from an extensive collection of 78 characteristics to enable precise classification, attaining remarkable accuracy over durations of 100 seconds, 200 seconds, and 300 seconds. The findings demonstrate that the respiration sensor is the most critical for the detection of stress. By utilizing supplementary information pertaining to individual conditions in various contexts, we can develop a framework to identify stress across multiple scenarios and accurately quantify stress levels, which can significantly aid healthcare professionals in prescribing suitable treatments. Looking ahead, future research should focus on refining machine learning algorithms for better predictive accuracy and exploring additional physiological indicators of stress. Moreover, the user experience must be prioritized to ensure that stress monitoring systems are accessible and engaging for diverse populations. Ultimately, the combination of machine learning and edge computing presents a robust framework for developing innovative solutions to address the growing challenge of stress, empowering individuals to lead healthier and more balanced lives. As this field continues to evolve, interdisciplinary collaboration among technologists, healthcare professionals, and researchers will be essential to realize the full potential of these technologies in improving mental health outcomes.

REFERENCES

[1] R. Alhalaseh and S. Alasasfeh, "Machine-learning-based emotion recognition system using EEG signals," *Computers*, vol. 9, no. 4, p. 95, 2020, doi: 10.3390/computers9040095.

[2] S. K. Bharti, S. Varadhaganapathy, R. K. Gupta, P. K. Shukla, M. Bouye, S. K. Hingaa, and A. Mahmoud, "Text-Based Emotion Recognition Using Deep Learning Approach," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, p. 2645381, 2022, doi: 10.1155/2022/2645381.

[3] M. B. Bin Heyat, F. Akhtar, S. J. Abbas, M. Al-Sarem, A. Alqarafi, A. Stalin, et al., "Wearable flexible electronics based cardiac electrode for researcher mental stress detection system using machine learning models on single lead electrocardiogram signal," *Biosensors*, vol. 12, no. 6, p. 427, 2022, doi: 10.3390/bios12060427.

[4] S. Gedam and S. Paul, "A review on mental stress detection using wearable sensors and machine learning techniques," *IEEE Access*, vol. 9, pp. 84045–84066, 2021, doi: 10.1109/ACCESS.2021.3085502.

[5] S. Ghosh, S. Kim, M. F. Ijaz, P. K. Singh, and M. Mahmud, "Classification of mental stress from wearable physiological sensors using image-encoding-based deep neural network," *Biosensors*, vol. 12, no. 12, p. 1153, 2022, doi: 10.3390/bios12121153.

[6] T. Iqbal, A. Elahi, W. Wijns, and A. Shahzad, "Exploring unsupervised machine learning classification methods for physiological stress detection," *Frontiers in Medical Technology*, vol. 4, p. 782756, 2022, doi: 10.3389/fmedt.2022.782756.

[7] L. Malviya and S. Mal, "A novel technique for stress detection from EEG signal using hybrid deep learning model," *Neural Computing and Applications*, vol. 34, no. 22, pp. 19819–19830, 2022, doi: 10.1007/s00521-022-07540-7.

[8] R. K. Nath, H. Thapliyal, A. Caban-Holt, and S. P. Mohanty, "Machine learning based solutions for real-time stress monitoring," *IEEE Consumer Electronics Magazine*, vol. 9, no. 5, pp. 34–41, 2020, doi: 10.1109/MCE.2020.2993427.

[9] T. Nijhawan, G. Attigeri, and T. Ananthakrishna, "Stress detection using natural language processing and machine learning over social interactions," *Journal of Big Data*, vol. 9, no. 1, p. 33, 2022, doi: 10.1186/s40537-022-00575-6.

[10] S. S. Panicker and P. Gayathri, "A survey of machine learning techniques in physiology based mental stress detection systems," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 2, pp. 444–469, 2019, doi: 10.1016/j.bbe.2019.01.004.

[11] A. I. Siam, S. A. Gamel, and F. M. Talaat, "Automatic stress detection in car drivers based on non-invasive physiological signals using machine learning techniques," *Neural Computing and Applications*, vol. 35, no. 17, pp. 12891–12904, 2023, doi: 10.1007/s00521-023-08428-w.

[12] O. Vargas-Lopez, C. A. Perez-Ramirez, M. Valtierra-Rodriguez, J. J. Yanez-Borjas, and J. P. Amezquita-Sanchez, "An explainable machine learning approach based on statistical indexes and SVM for stress detection in automobile drivers using electromyographic signals," *Sensors*, vol. 21, no. 9, p. 3155, 2021, doi: 10.3390/s21093155.

[13] R. Walambe, P. Nayak, A. Bhardwaj, and K. Kotecha, "[Retracted] Employing Multimodal Machine Learning for Stress Detection," *Journal of Healthcare Engineering*, vol. 2021, no. 1, p. 9356452, 2021, doi: 10.1155/2021/9356452.

[14] H. Zhang, L. Feng, N. Li, Z. Jin, and L. Cao, "Video-based stress detection through deep learning," *Sensors*, vol. 20, no. 19, p. 5552, 2020, doi: 10.3390/s20195552.

[15] L. Zhu, P. Spachos, P. C. Ng, Y. Yu, Y. Wang, K. Plataniotis, and D. Hatzinakos, "Stress detection through wrist-based electrodermal activity monitoring and machine learning," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 5, pp. 2155–2165, 2023, doi: 10.1109/JBHI.2023.3239305.