J. Electrical Systems 20-7s (2024): 2143-2150

[1]Vanisha P. Vaidya
[2]Dr. Suresh S. Asole

Mental Stress Prediction Using Machine Learning on Real Time Dataset

Journal of Electrical Systems

Abstract: - Mental stress is a significant issue affecting individuals across all age groups in today's society. The impact of stress extends beyond mental well-being and can lead to various chronic illnesses, including depression, malignancy, and cardiovascular disease (CVD). Addressing mental stress and implementing strategies for stress management and prevention are crucial not only for mental well-being but also for reducing the risk of associated chronic illnesses. In this paper, we are focussing of the problem of mental stress in today's society and the significance of early prediction and management. The use of Random Forest (RF) method with enhanced band pass filtration technique on ECG data for predicting stress levels is a promising approach. Achieving a 96.73% accuracy in stress categorization, along with improvements over prior research results, highlights the effectiveness of the proposed model. The main contribution of the research work is that the results of the model are validated using real-time datasets which further strengthens its reliability and applicability in practical scenarios.

Keywords: Mental Stress, Random Forest, ECG Signals, Real time dataset

1. Introduction:

Stress typically arises when an individual's mental and physiological capacities are insufficient to meet the demands placed upon them [1]. This can occur when an individual is confronted with unfamiliar material or when a student is faced with multiple unresolved problems within a limited timeframe for a final examination. Stress is a pervasive phenomenon observed in contemporary society, exerting a significant impact on the everyday lives and overall well-being of individuals. Based on data from the American Physiological Association and American Institute of Stress, in 2014, 77% of individuals in the United States consistently had physical symptoms, while 73% experienced psychological symptoms. These symptoms were attributed to stress [2]. In addition, 33% of individuals reported experiencing severe stress, while 48% of individuals acknowledged that stress has detrimental effects on both their personal and professional spheres. In the EU, a comparable situation occurred, with over 22% of employees perceiving their health as being jeopardised due to their demanding work [3]. The ensuing expenses incurred due to stress-induced healthcare and absenteeism were incalculable.

In the United States, it amounted to a staggering 300 billion USD annually[2]. In the United Kingdom, an annual loss of 13 million working days occurred, resulting in a financial burden of 12 billion pounds [4]. Therefore, in order to alleviate stress and maintain a state of well-being, the identification and control of stress have become imperative and indispensable for individuals of average means. The autonomic nervous system (ANS) would initially exhibit the physiological response to stress[5]. The ANS is composed of two branches, namely the sympathetic and parasympathetic divisions. The occurrence of stress events has the potential to disrupt the equilibrium between the two branches. The suppression of the parasympathetic branch would result in a hyper-activation of the sympathetic branch. The transmission of this information would be directed towards the heart activity, which may be quantified by the analysis of electrocardiogram (ECG) data. Therefore, the electrocardiogram (ECG) signal is commonly utilised as a primary indication for stress assessment. Heart rate variability (HRV) refers to the fluctuation in the time interval between two successive R peaks (heartbeats) of electrocardiogram (ECG) signals [6]. HRV is considered resilient to noise and disturbances due to its reliance solely on R peaks for computation, as these peaks exhibit the largest amplitude in

[1]vanisha.vaidya@gmail.com
Department of Computer Science and Engineering
Dr. A. P. J. Abdul Kalam University, Indore ,India
[2]suresh_asole@yahoo.com
Department of Computer Science and Engineering
Babasaheb Naik College of Engineering,Pusad ,India

electrocardiogram (ECG) signals. Therefore, heart rate variability (HRV) is frequently employed as a diagnostic instrument for stress detection. In order to get a heart rate variability (HRV) measurement, it is often necessary to connect three electrocardiogram (ECG) electrodes to the right leg (RL), right arm (RA), and left leg (LL). These electrodes are used to capture the standard Lead II ECG signal[7]. After the identification of all the R peaks, it is possible to extract HRV characteristics for the purpose of stress detection [8]. Typically, a certain time period is set, and the characteristics are computed during that period. The research on stress detection using HRV typically use a window length measured in minutes [9],[10]. Therefore, the ability to make inferences about cognitive stress is limited to a minimum delay of one minute, which is not feasible for real-time stress detection.

Numerous techniques have been suggested for the purpose of feature extraction, drawing upon physiological manifestations of stress such as heightened heart rate[9], alterations in the balance between the sympathetic and parasympathetic branches of the autonomic nervous system [6], and irregularities in heart rate rhythm [11], among others. Stress detection commonly involves measuring heart rate (HR) and calculating the power ratio between the low-frequency band (0.04 - 0.15 Hz) and the high-frequency band (0.15 _ 0.4 Hz), among other factors. Nevertheless, various investigations have indicated that the efficacy of a solitary feature is not uniform due to significant inter-participant variations. Consequently, the integration of multiple characteristics has been proven to yield superior performance in stress detection [10]. This finding suggests that individuals' reactions to stress might differ and have various expressions in heart rate variability (HRV). Subsequently, the distinguished characteristics were transmitted to the classifier for identification. Linear discriminant analysis (LDA) and support vector machine (SVM) are two often used traditional classifiers [12].

The convolutional neural network (CNN) is a type of artificial neural network (ANN) that commonly includes a convolutional layer inside its hidden layers [13]. The technique was first utilised in computer vision applications and has since garnered attention in various fields, including biosignal classification[14],[15]. For instance, it has been employed in the classification of electromyogram (EMG) signals for gesture recognition[16] and the identification of electroencephalogram (EEG) patterns for assistive machine control [14]. CNN has been effectively utilised in several applications such as arrhythmia detection[17], signal component identification [18], and biometric recognition [19] in the field of ECG signal categorization. These experiments showcased the efficacy of CNN in classifying bio signals, suggesting the potential of CNN in detecting stress based on ECG data.

To the authors' knowledge, this work is the first to utilise standard machine learning with the Random Forest approach for real-time detection of mental stress using ECG data. Specifically, it aims to infer mental stress from specific segments of ECG signals. Real-time stress detection has become a crucial requirement in several practical applications, particularly in the context of acute cognitive stress detection. This necessitates the ability to make prompt judgements.

The primary objective of this work was to investigate the identification of acute cognitive stress utilising the RF technique on ECG signals within a real-time dataset. We conducted a comparative analysis of the performance of the suggested RF and standard ECG based stress estimation techniques. Our findings demonstrate that our proposed approach outperformed the old methods in terms of accurately predicting mental stress on real-time datasets. The objective of this work was to enhance the practical implementation of electrocardiogram (ECG)-based stress measurement by offering expedited, dependable, and instantaneous outcomes.

2. Methodologies

In this part, we will go over the recommended techniques. Figure 2.1 depicts the total system architecture diagram.

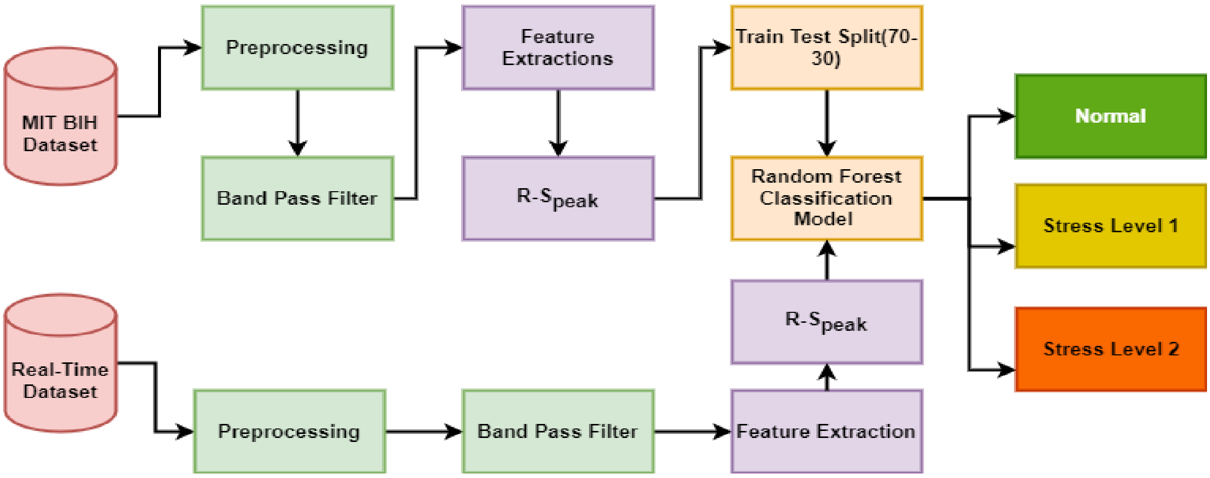

Figure 2.1: System Architecture

2.1 Dataset Description:

2.1.1 Training Dataset:

The proposed study uses MIT-BIH Database for training, where ECG recording of 47 patients were taken. Twenty-three recordings were chosen at random from a set of 4000 24-hour mobile ECG recordings gathered at Boston's Beth Israel Hospital from a variety of groups of patients (approximately 60%) and outpatients (approximately 40%). Over a 10-mV range, the video recordings were digitized at 360 bits per second per lane with 11-bit resolution. Each record was annotated individually by two or more cardiologists; disagreements were resolved in order to get the computer-readable references annotations for each heartbeat (about 110,000 evaluations in total) included in the database.

2.1.2 TESTING DATASET:

As we need to work on real-time dataset, we contacted various cardiologist for getting the real time ECG signal data for stress classification and prediction. We Collected the dataset from one of the famous cardiologists from Mumbai, India region. They provided us the ECG signal dataset of 11 patients for real time stress analysis for our research work.All this 11 patient's ECG signal dataset cannot be shown here cause of paper length

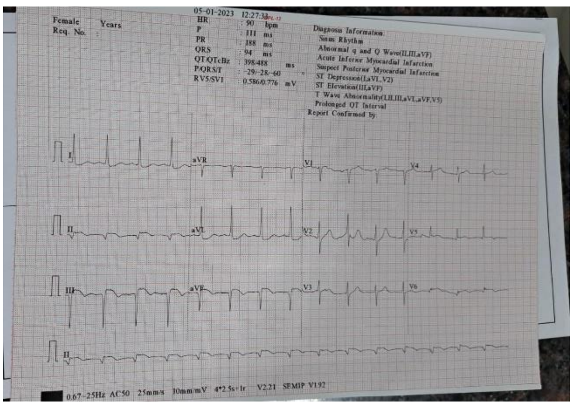

Figure 2.1: Patient 1 ECG Signal

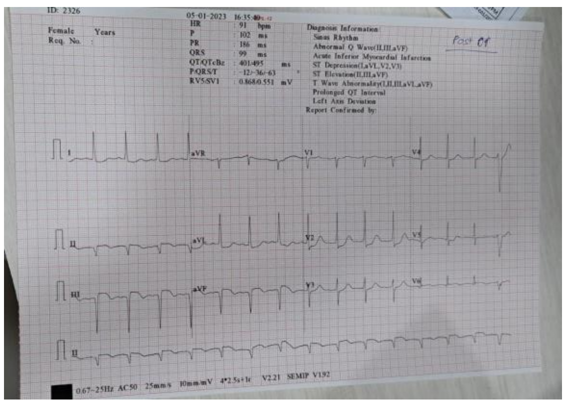

Figure 2.2: Patient 2 ECG Signal

constraint, but 2 patients' data was shown in following figures 2.1 and 2.2.

2.2 Preprocessing And Feature Extractions: Electrocardiography is a non-invasive way of assessing a person's health that examines the electrical status of the heart. Several variables cause noise during an EKG, severely lowering ECG classification accuracy. We used a band-pass filter to solve this problem and discovered that a band-pass filter has a sample rate of 360 Hz & a cutoff frequency of 150 Hz removed 90.89% of the noise.

Figure 2.2 shows R Speak values obtained from an ECG wave. dividing these data during and without stress allows for a more reliable examination of the ECG.Rpeak&Speak were derived from an electrocardiogram (ECG) after a threshold was applied. Rpeak extracted the pole if the threshold level was greater than or equal to 0.2mV in one time frame of the signal; if the criteria value were less than 0.54mV in one time frame, Rpeak extracted the pole. The heart beats irregularly and quickly during stress, the gap between the R and R parts of the ECG signal narrows, while the R Speak increases.In the unstressed condition, on the other side of the hand, the heart is generally steady, the R-R gap expands, and the R Speak lowers. The standard deviation (SD) of R Speak with stress was 1.47 mV in each test, and 4.25 mV when stressed.

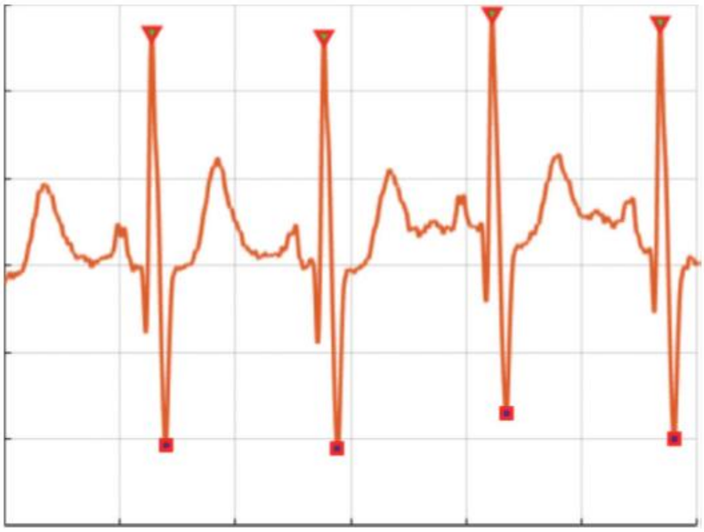

Figure 2.2: Feature Extraction by Threshold Values

2.3 Model Designing:

We utilized most of the machine learning classification model for categorization of stress and from the result analysis of performance of each model, we selected the best model as Random Forest based on key performance metrics such as categorization accuracy, precision, recall, and F1-Score.

2.3.1 Random Forest:

To increase prediction accuracy, a bagging approach known as Random Forest is employed to blend several decision trees. Individuals are taught bagging on their own. In this technique, numerous data samples are constructed from a single dataset using replacement, and each of the decision trees is trained on a different set of data samples. The tree's characteristics are also picked at random throughout the construction process. A majority vote can be used to aggregate the forecasts of numerous trees. Increasing the random forest's accuracy through optimizing variables such as the number of estimation tools, the smallest possible size of the node, and the number of attributes used to partition nodes

3. Experimental Results:

After applying the random forest model for classification of stress on 11 patients real time dataset, we found the analysis of the patient’s records based on ECG signals of 30 sec ultrashort wavelength into the 3 classification stages as Normal (Stress Level 0), Moderate (Stress Level 1) and High (Stress Level 2). The results are discussed in table 3.1.

Table 3.1: Real Time Data Stress Prediction

	BPM	P(s)	PR(s)	QRS(s)	Execution Time	Prediction
Patient1	90	0.111	0.188	0.04	0.17	Moderate
Patient2	91	0.102	0.186	0.099	0.21	Moderate
Patient3	86	0.115	0.185	0.089	0.19	Moderate
Patient4	64	0.101	0.159	0.092	0.14	Moderate
Patient5	48	0.116	0.18	0.09	0.28	High
Patient6	64	0.094	0.198	0.093	0.15	Normal
Patient7	86	0.081	0.152	0.067	0.21	Moderate
Patient8	106	0.072	0.151	-0.05	0.2	High
Patient9	103	0.101	0.145	0.103	0.277	Moderate
Patient10	68	0.094	0.157	0.084	0.207	Moderate
Patient11	59	0.107	0.164	0.144	0.2477	High

Out of 11 patient's ECG prediction for stress, we found only 1 patient as normal,3 patients with high stress and rest 7 patients having moderate level of stress. The Random Forest model has produced 96.73% accuracy for the stress classification and prediction when tested on real time dataset which is higher comparative with other studies with real time classification of the stress using ECG Signals.

Figure 3.1 show the stress prediction pie chart for real time dataset of 11 patients. We found that 64% patients are under moderate stress, 27% are under high stress and 9% are under normal stress.

When we done the comparative analysis with our research work with most recent research work then we found

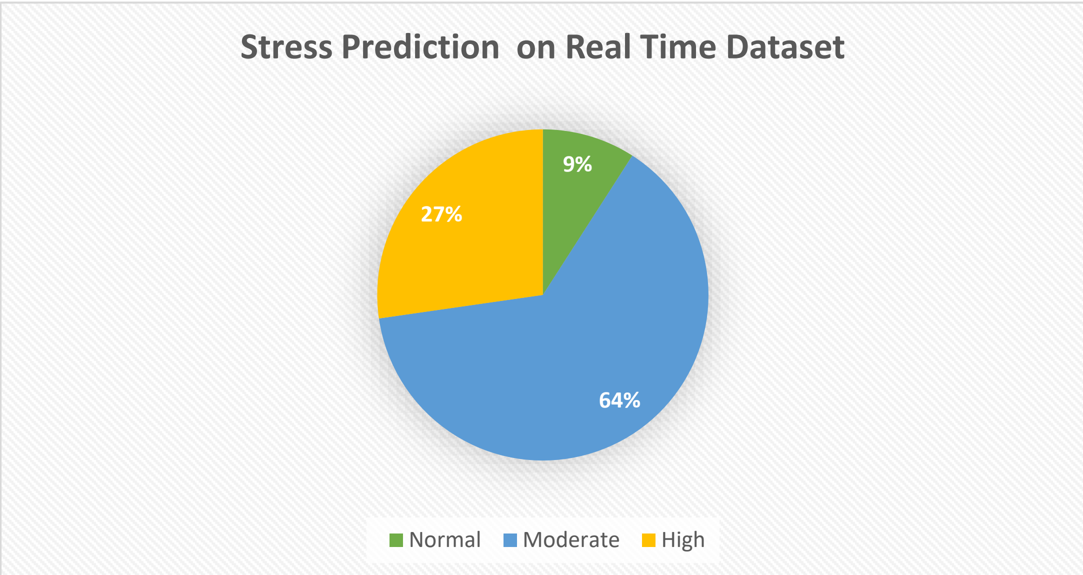

Figure 3.1: Stress prediction pie chart for real time dataset

that our model has performed better compared to all the recent research work. We are comparing on the basis of their model's performance on ECG dataset for stress classification. Following table 3.2 summarizes the comparative analysis with most recent work.

Table 3.1: Comparative Analysis of the proposed work

Reference	Accuracy	Model	Type of Signals	Window Size	Stress Classification
[20]	90.19	CNN	ECG	10 s	2
[21]	89.8	CNN	ECG	60 s	2
[22]	87.39	CNN-RNN	ECG	10 s	2
[23]	83.9	CNN	ECG and RSP	50 s	2
[24]	82.7	CNN	ECG	10 s	2
[25]	92.8	CNN	ECG	25 s	3
[26]	86.5	CNN-BiLSTM	ECG	10 s	3
[27]	85.45	CNN	ECG	30 s	3
[28]	80.00	SVC	ECG	60s	2
[29]	98.5	KNN	ECG	60s	2
Proposed Work	96.73	Random Forest	ECG	30s	3

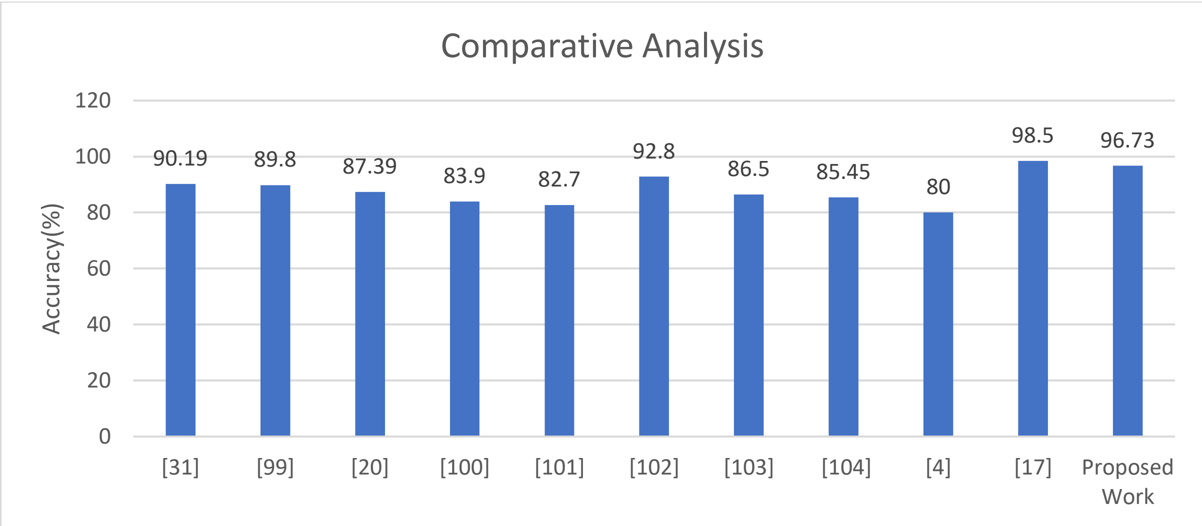

Figure 3.2: Comparative Analysis

4. Conclusion: By applying the random forest which was best performing model during the training our database, on the real time database procured by the cardiologist from Mumbai, the RF model produced 96.73% accuracy. When tested with ECG signal database of 11 patients, we found that 1 patient got into the normal stress classification,7 into Moderate stress and 3 into high stress classification. This is one of the unique studies as per our knowledge who applied Stress classification with Random Forest classification algorithm with 96.73% accuracy on real time database with early prediction of average prediction time as 0.17 seconds.

Acknowledgement: We sincerely thanks to Dr. Kapil Rathi (Consultant Physician, Cardiac and Chest Specialist) from My Health Clinic, Andheri(E), Mumbai, India for providing the real time dataset of ECG Signals.

REFERENCES:

[1] Y. M. Ulrich-Lai and J. P. Herman, “Neural regulation of endocrine and autonomic stress responses,” *Nat. Rev. Neurosci.*, vol. 10, no. 6, pp. 397–409, 2009, doi: 10.1038/nrn2647.

[2] American Psychological Association. (2015), “Stress in America: Paying With Our Health,” *Am. Psychol. Assoc.*, pp. 1–61, 2018, [Online]. Available: https://www.apa.org/news/press/releases/stress/2014/stress-report.pdf

[3] “Evaluation of an Integrated System of Wearable Physiological Sensors for Stress Monitoring in Working Environments by Using Biological Markers,” *IEEE Trans. Biomed. Eng.*, vol. 65, no. 8, pp.

1748–1758, 2018, doi: 10.1109/tbme.2017.2764507.

[4] "Stress at work: a report prepared for The Work Foundation's principal partners : WestminsterResearch." https://westminsterresearch.westminster.ac.uk/item/91wvv/stress-at-work-a-report-prepared-for-the-work-foundation-s-principal-partners (accessed Mar. 23, 2024).

[5] G. E. Billman, H. V. Huikuri, J. Sacha, and K. Trimmel, "An introduction to heart rate variability: Methodological considerations and clinical applications," *Front. Physiol.*, vol. 6, no. FEB, 2015, doi: 10.3389/FPHYS.2015.00055.

[6] F. Lombardi and P. K. Stein, "Origin of Heart Rate Variability and Turbulence: An Appraisal of Autonomic Modulation of Cardiovascular Function," *Front. Physiol.*, vol. 2, 2011, doi: 10.3389/fphys.2011.00095.

[7] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on Psychological Stress Detection Using Biosignals," *IEEE Trans. Affect. Comput.*, vol. 13, no. 1, pp. 440–460, 2022, doi: 10.1109/TAFFC.2019.2927337.

[8] R. Castaldo, P. Melillo, U. Bracale, M. Caserta, M. Triassi, and L. Pecchia, "Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis," *Biomed. Signal Process. Control*, vol. 18, pp. 370–377, 2015, doi: 10.1016/j.bspc.2015.02.012.

[9] D. McDuff, S. Gontarek, R. P.-2014 36th annual, and undefined 2014, "Remote measurement of cognitive stress via heart rate variability," *ieeexplore.ieee.orgD McDuff, S Gontarek, R Picard2014 36th Annu. Int. Conf. IEEE Eng. 2014•ieeexplore.ieee.org*, pp. 2957–2960, 2014, doi: 10.1109/EMBC.2014.6944243.

[10] R. Ribeiro and J. P. Cunha, "A regression approach based on separability maximization for modeling a continuous-valued stress index from electrocardiogram data," *Biomed. Signal Process. Control*, vol. 46, pp. 33–45, Sep. 2018, doi: 10.1016/j.bspc.2018.06.006.

[11] R. McCraty and M. A. Zayas, "Cardiac coherence, self-regulation, autonomic stability and psychosocial well-being," *Front. Psychol.*, vol. 5, no. SEP, 2014, doi: 10.3389/FPSYG.2014.01090.

[12] M. Kusserow, O. Amft, and G. Troster, "Modeling arousal phases in daily living using wearable sensors," *IEEE Trans. Affect. Comput.*, vol. 4, no. 1, pp. 93–105, 2013, doi: 10.1109/t-affc.2012.37.

[13] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, no. 3–4, pp. 197–387, 2013, doi: 10.1561/2000000039.

[14] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *J. Neural Eng.*, vol. 14, no. 1, p. 16003, 2016, doi: 10.1088/1741-2560/14/1/016003.

[15] X. Zhai, B. Jelfs, R. H. M. Chan, and C. Tin, "Self-Recalibrating Surface EMG Pattern Recognition for Neuroprosthesis Control Based on Convolutional Neural Network," *Front. Neurosci.*, vol. 11, 2017, doi: 10.3389/fnins.2017.00379.

[16] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, and J. Li, "Gesture recognition by instantaneous surface EMG images," *Sci. Rep.*, vol. 6, no. 1, 2016, doi: 10.1038/srep36571.

[17] J. Torres Ruiz, J. D. Buldain Pérez, and J. R. Beltrán Blázquez, "Arrhythmia Detection Using Convolutional Neural Models," *Jorn. Jóvenes Investig. del I3A*, vol. 7, 2019, doi: 10.26754/jji-i3a.003522.

[18] Y. Xiang, Z. Lin, and J. Meng, "Automatic QRS complex detection using two-level convolutional neural network," *Biomed. Eng. Online*, vol. 17, no. 1, 2018, doi: 10.1186/s12938-018-0441-4.

[19] R. Donida Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti, "Deep-ECG: Convolutional Neural Networks for ECG biometric recognition," *Pattern Recognit. Lett.*, vol. 126, pp. 78–85, 2019, doi: 10.1016/j.patrec.2018.03.028.

[20] H.-M. Cho, H. Park, S.-Y. Dong, and I. Youn, "Ambulatory and Laboratory Stress Detection Based on Raw Electrocardiogram Signals Using a Convolutional Neural Network," *Sensors* , vol. 19, no. 20. 2019. doi: 10.3390/s19204408.

[21] G. Giannakakis, E. Trivizakis, M. Tsiknakis, and K. Marias, "A novel multi-kernel 1D convolutional neural network for stress recognition from ECG," *2019 8th Int. Conf. Affect. Comput. Intell. Interact. Work. Demos, ACIIW 2019*, pp. 273–276, 2019, doi: 10.1109/ACIIW.2019.8925020.

[22] H. B, Y. J, V. T, M.-G. I, P. C, and Z. BT, "Deep ECGNet: An Optimal Deep Learning Framework for Monitoring Mental Stress Using Ultra Short-Term ECG Signals," *Telemed. J. E. Health.*, vol. 24, no.

10, pp. 753–772, Oct. 2018, doi: 10.1089/TMJ.2017.0250.

[23] W. Seo, N. Kim, S. Kim, C. Lee, and S. M. Park, "Deep ECG-respiration network (DeepER net) for recognizing mental stress," *Sensors (Switzerland)*, vol. 19, no. 13, pp. 1–16, 2019, doi: 10.3390/s19133021.

[24] J. He, K. Li, X. Liao, P. Zhang, and N. Jiang, "Real-Time Detection of Acute Cognitive Stress Using a Convolutional Neural Network from Electrocardiographic Signal," *IEEE Access*, vol. 7, pp. 42710–42717, 2019, doi: 10.1109/ACCESS.2019.2907076.

[25] J. Huang, X. Luo, and X. Peng, "A novel classification method for a driver's cognitive stress level by transferring interbeat intervals of the ECG signal to pictures," *Sensors (Switzerland)*, vol. 20, no. 5, 2020, doi: 10.3390/s20051340.

[26] P. Zhang *et al.*, "Real-time psychological stress detection according to ECG using deep learning," *Appl. Sci.*, vol. 11, no. 9, 2021, doi: 10.3390/app11093838.

[27] Z. Ahmad and N. M. Khan, "Multi-level Stress Assessment Using Multi-domain Fusion of ECG Signal," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2020-July, pp. 4518–4521, 2020, doi: 10.1109/EMBC44109.2020.9176590.

[28] M. Chauhan, S. V. Vora, and D. Dabhi, "Effective stress detection using physiological parameters," *Proc. 2017 Int. Conf. Innov. Information, Embed. Commun. Syst. ICIIECS 2017*, vol. 2018-Janua, pp. 1–6, 2017, doi: 10.1109/ICIIECS.2017.8275853.

[29] U. R. Acharya *et al.*, "Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ECG signals: A comparative study," *Inf. Sci. (Ny).*, vol. 377, pp. 17–29, Jan. 2017, doi: 10.1016/j.ins.2016.10.013.