High-Accuracy Stress Detection Using Wrist-Worn PPG Sensors

Anice Jahanjoo*, Nima TaheriNejad+*, Amin Aminifar+
*Technische Universität Wien (TU Wien), Austria, +Heidelberg University, Germany
anice.jahanjoo@tuwien.ac.at, {nima.taherinejad, amin.aminifar}@ziti.uni-heidelberg.de

***Abstract*—Stress has become a prevalent issue affecting individuals' physical and mental well-being. Detecting stress is the first crucial step to managing it and preventing it from causing other health issues. In this paper, we present a new method to improve the performance of detecting stress, using a comfortable to wear sensor, namely Photoplethysmography (PPG), which is embedded virtually in all smartwatches. To this end, we use PPG sensor data from the publicly available wearable stress and affect detection dataset (WESAD). Using new denoising processes, segmentation methods, and key feature extract, we achieve 95.55% accuracy in detecting stress using the Support Vector Machine (SVM) algorithm. Simplifying the process alongside improved accuracy in this paper facilitates smartphone usage as a real-time stress detection, which we plan as future work.**

***Index Terms*—Stress detection, machine learning, wearable sensors, photoplethysmography, denoising, segmentation, feature extraction.**

I. Introduction

Stress is an individualized experience that varies among people based on their unique susceptibility and resilience, as well as the specific tasks they engage in [1]. It has a major influence on human behavior, performance during work, personal relationships, and bodily health [2]. Generally, there are two types of stress that can potentially contribute to the development of different diseases that are chronic and acute. At the beginning of the "Major Depressive Episode" research, researchers found there is a significant link between chronic and acute stress [3]. Furthermore, persistent stress was also connected to the onset of sudden events, and there was a tendency indicating that a higher level of acute stress is more strongly correlated with depression in individuals experiencing high chronic stress compared to those with low chronic stress [3]. In addition to depression as an impact of acute stress on individuals' health, chronic stress exerts a notable impact on the immune system [4], increasing heart attacks and strokes, and eventually leading to the development of various illnesses [5].

In recent years, with the advancement of technology, many different sensors have been available to monitor the physiological states of individuals [6]–[11]. Sensors could be invasive (which need to be placed under the skin or attached to a special part of a body), which makes it difficult to be used by the general public. The other type of sensors are non-invasive and could be embedded into a wristband, headband, or even a ring to be attached to human skin, which makes them convenient [12], [13]. PPG, as a sensor that could be used in a wearable device, is an optical and non-invasive sensor that quantifies alterations in skin hue linked to changes in blood volume within the subcutaneous blood vessels throughout the cardiac cycle. PPG sensors employ optical pulses produced by a light source and capture the reflected light using a photodetector [14]. Fig. 1 displays a segment of a Blood Volume Pulse (BVP) signal, which is the phasic change in blood volume that corresponds to each heartbeat interval and obtained by wearing a PPG sensor.

This paper aims to improve the performance of detecting stress using PPG sensor data. To this end, we propose applying denoising methods and extracting a novel combination of only seven features. We demonstrate that these techniques enhance performance when compared to previous studies that utilized a larger set of features. We also investigate different machine learning algorithms to classify the stress in subjects. To evaluate our proposed algorithm, we use WESAD [15] which is a publicly available dataset collected in a laboratory.

We provide an overview of the existing literature related to stress detection in Section II, and an overview of our proposed methods and techniques for each step of processing PPG signal to detect stress in Section III. Section IV provides findings based on our algorithm and compares them with previous studies. In the end, Section V concludes the paper and discusses future works.

II. Background

In recent years, there has been a growing interest in utilizing sensors to investigate the human body using various process-

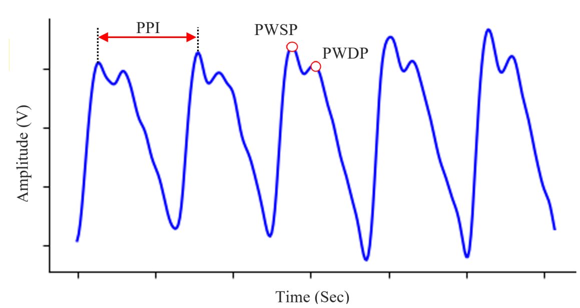

Fig. 1. BVP signal (PPI: Peak to Peak Interval, PWSP: Pulse Wave Systolic Peak, PWDP: Pulse Wave Diastolic Peak)

This work has been partially funded by the Vienna Science and Technology Fund (WWTF) [10.47379/ICT20034].

ing techniques [16]–[22]. Different sensor data can provide valuable vision into individuals' mental states, including their emotions and overall well-being [7], [23]–[25] or their physical states, particularly in the case of vulnerable individuals [26], [27]. Among these sensors, PPG is an attractive option for researchers aiming to detect mental states, specifically stress. This can be done either in combination with other physiological sensors or as a standalone tool. To detect stress, researchers have used traditional and modern feature extraction and machine learning algorithms. Das *et al.* [28] used different sensors, including PPG, Electroencephalogram (EEG), and Galvanic Skin Response (GSR), to detect stress in their work. They preprocessed signals with a Band-Pass (BP) filter and extracted 18 features from GSR, 62 features from PPG, and 1476 features from EEG. Unfortunately, they did not report the window length they used for the PPG data. Using a clustering method, they achieved 69% F1-Score for detecting stress on a ten-subject dataset. Despite using an extensive set of features, the results were not satisfactory.

Hasanpoor *et al.* [29] utilized a deep learning model to detect stress without employing any denoising techniques. Instead, they relied on the CNN-MLP algorithm, which demonstrated the capability to identify motion artifact noises. Additionally, their approach did not involve a segmentation step, and the raw signal data served as the input for the CNN-MLP model. As a result, they achieved an accuracy of 82% in detecting stress instances from non-stress instances. The employment of complex models like Convolutional Neural Network (CNN) led to significant computational overheads that proved unnecessary, given that better results can be achieved through traditional algorithms. Benchekroun *et al.* [30] Compared the performance of Electrocardiograms (ECG) and PPG sensor data to detect stress. They used BP and Notch filtering and a 5-minute window for segmentation. Random Forest (RF) is used for stress classification in their study on samples with 22 features that involved frequency and time domain features. They divided their dataset with 80% for training and 20% for the test, which means there were data of a same subject in both train and test sets. Ultimately, they could detect stress with 83% accuracy by using PPG sensor data. Including data from the same subject in both the train and test sets can limit the model's generalization ability, potentially reducing its accuracy when encountering data that was not present in the training set.

III. Proposed Methods

A. Algorithm Overview

Fig. 2 shows the proposed method that we use in this work for stress detection. First, to remove artifacts and noises, we start with filtering noise from the PPG signal. To extract the features of interest, we divide the signal into 360-second windows with 30-second sliding windows. As opposed to commonly used rectangular windowing, we use Hanning windows.

Features that we extract from each window are Time-Domain, Frequency-Domain, and Non-Linear-Domain. From a

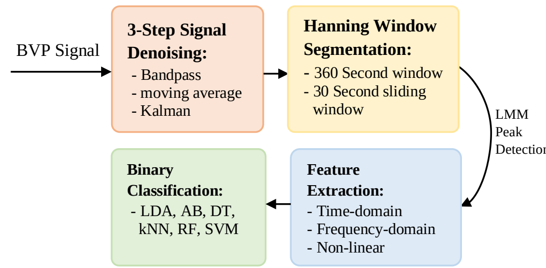

Fig. 2. Overall process of the proposed method

comprehensive list of features commonly employed in state-of-the-art [15], [30], [31], we have carefully selected a smaller, but vital set of features across various domains. This selection aims to reduce computational complexity while achieving superior results. In the end, binary classifications are done using some basic binary classifiers to detect stress samples and compare results with other research.

B. Noise Filtering

One of the most challenging parts in processing BVP is noise. Noise in this type of signal could have different reasons, and one of them is body movement and activities, which are unavoidable, especially in real-life data collection. In this paper, to reduce the effects of the motion artifacts and filter the noise, we use BP with a low-pass cutoff frequency of 0.5 and a high-pass cutoff frequency of 10 and Kalman filtering [32]. As the final step of denoising and smoothing the signal, a 3-point Moving Average (MA) is applied to the signal.

C. Segmentation

In order to detect stress, it is necessary to understand the characteristics of BVP signal. We compute Heart Rate Variability (HRV) as the primary parameter for detecting stress. In order to compute HRV parameters, the signal should be segmented to some windows with the appropriate sizes.

To assess how the window length influences the accuracy of stress detection, we isolated the continuous data segments labeled as stress for each subject in the dataset. We then examined the segmentation using various window sizes. In this research, a 360-second window is chosen because of its better performance than windows with shorter lengths. Fig. 3 shows how much window size could change the accuracy of detecting stress in the dataset. Exploring window lengths longer than 360 seconds is not feasible due to the length of signals within the dataset.

To avoid losing peaks and other important data between each two windows, a 30-second sliding window is considered in computing windows. In contrast to the state of the art, where rectangular windows are used, in this paper, we use Hanning windows [33] to achieve enhanced outcomes, in particular for frequency-domain features. Hanning windowing is considered better than rectangular windowing because it helps reduce spectral leakage, which is the leakage of power from one frequency to another in the frequency domain. By smoothing

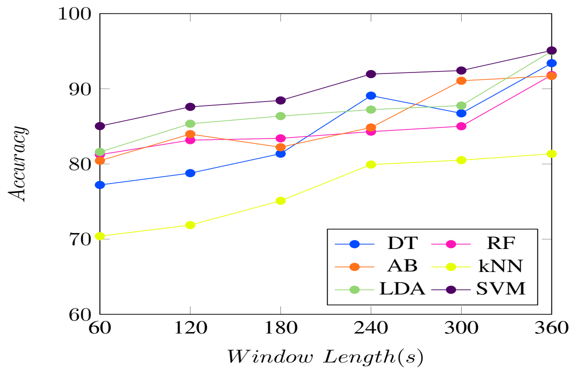

Fig. 3. Window length impact on stress detection accuracy for different algorithms

the edges of the window, it reduces the abrupt changes that can introduce unwanted artifacts and distortion in the frequency analysis. This smoother transition helps to provide more accurate and reliable frequency domain representations of the signal. To detect peaks and utilize them for extracting features, we use Local Maxima Method (LMM) [34] to retrieve all the local maximum points within the signal. After obtaining these points, we remove those with a value lower than the average amplitude of the entire signal.

D. Feature Extraction

One commonly utilized feature for stress detection is HRV. HRV refers to the variation in the time between heartbeats (NN) or the peaks mentioned above. Stress typically leads to decreased HRV, indicating the dominance of sympathetic nervous system activity. In [31] authors present 29 common HRV features that are important to find information from BVP signals. HRV features such as mean time between heartbeats (MNN), the ratio of adjacent NN intervals differing by more than 50ms (pNN50), or total power, are computed to quantify the variability of the heart rate [31]. In this study, one of our objectives is to enhance results while employing a reduced set of features. Consequently, we explore commonly used HRV features and empirically identify the seven essential features, including HRV time-domain, frequency-domain, and nonlinear-domain components, to achieve high accuracy in stress detection. These features are described in Table I.

E. Classification Algorithms

The seven features described earlier are utilized as input for the classification process. We applied and compared six machine learning algorithms to detect stress samples.

TABLE I
HRV Features Used for Stress Detection

Domain	Feature Name	Feature Description
Time	HRM	Mean of the heart rates
	HRSD	Standard deviation of the HRs
	MNN	NN intervals mean
	pNN50	% of absolute differences in successive NN values >50 ms
Frequency	Total Power	Sum of the energy in the ULF, VLF, LF, and HF bands
Non-linear	Approximate Entropy	Segment regularity and complexity
	D2	Correlation dimension

k-Nearest Neighbours (kNN), Linear Discriminant Analysis (LDA), SVM, Decision Tree (DT), RF, and AdaBoost (AB) are binary classifiers that we used in this paper. To implement the entire data processing and make it comparable with other works, we utilized the scikit-learn implementation [35] of these classifiers. For kNN classifier, based on our experience on this dataset, we found nine as a number of neighbors to achieve the highest accuracy in classifying stress data. In the case of the AB ensemble learner, the decision tree function was utilized as the base estimator. Within each of the decision-tree-based classification algorithms, information gain was employed to evaluate the effectiveness of splitting decision nodes and a minimum sample threshold of 20 was applied for node splitting. We configured RF and AB to utilize 100 base estimators.

IV. Results and Comparisons

A. Setup

In this research, we utilize the PPG sensor dataset from the publicly available WESAD dataset [15]. The dataset provides multi-sensor data collected from 15 subjects using Empatica E4 and RespiBAN Professional devices during a lab study.

For the purpose of this paper, which is to achieve optimal results with minimal computation, we focus exclusively on the PPG sensor data, which captures BVP data and is particularly important for stress detection. The data recorded by the Empatica E4 wrist-worn device were collected at 64Hz sampling rate.

To create a division between the training and testing sets, we allocate 14 subjects' generated samples for training machine learning algorithms, while the remaining subjects' data is reserved as the test set. In each step, we use 207 samples as train samples and 15 samples as test samples. We repeat this process for each subject, and the results presented in the tables are the average evaluation of all 15 subjects' data. This cross-validation approach ensures that the algorithms are never tested on the same data they were trained on.

B. Evaluation Metrics

To evaluate the performance of our predictions, we consider a measure of all the correctly identified samples alongside the measure of the incorrectly classified samples. Consequently, we use Accuracy (ACC) metric, which is calculated using True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) to assess performance based on stress and non-stress samples that our algorithms classified correctly. We measure F1-Score (F1) metric, which is calculated using precision and recall [36] to show the performance of the proposed method as a better measure of the incorrectly classified samples. Also, Area Under the Curve (AUC) [37] was used to measure the proposed method's performance. The evaluation metrics described in this section are expressed as follows:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{2}$$

C. Results

In Table II, we present the effectiveness of each filtering method in segmentation with a 360-second Hanning window and a 30-second sliding window. We obtained 222 windows and computed proposed features to generate 222 samples, including 161 non-stress samples and 61 stress samples as inputs for the classifiers. To identify the optimal features, we began with a set of 24 features from F. Shaffer *et al.* [31]. Through a 6-step process, we selected features with minimal correlation. At each step, we used a set of features that includes all three categories of time and frequency domain and nonlinear features. Fig. 4 demonstrates that for this dataset, reducing the number of features until identifying the essential features not only maintains accuracy but can also enhance it for most of the algorithms. In this part we empirically found seven important that provide the highest accuracy. A simplified model with fewer features is often more robust and less sensitive to variations and changes in the data, leading to improved accuracy across different datasets.

The reported results in Table II show that adding each step improved the accuracy of most classification algorithms. Confirming the effectiveness of the proposed method. Among the six classifiers that we used to evaluate the performance of the proposed method, the highest accuracy, F1-Score and AUC is achieved by SVM, which detects stress with 95.55% accuracy and has a 91.42% F1-Score. SVM aims to find an optimal hyperplane that separates data points of different classes, which can lead to better stress detection results by effectively classifying stress-related data. This shows our proposed method's good performance in both evaluation aspects: detecting stress correctly and ignoring non-stress samples.

D. Comparison Results

To ensure a fair comparison of our proposed method with others, it is important to evaluate it against studies that use a similar setup, i.e., the same dataset, sensors, and algorithms. We compared our results with the work of P. Schmit *et al.* [15], which had similar conditions but used different preprocessing methods. The results of their study are presented in Table III.

TABLE II
PERFORMANCE COMPARISON FOR USING DIFFERENT DENOISING AND SEGMENTATION METHODS ON DIFFERENT BINARY (STRESS, NON-STRESS) CLASSIFIERS. EVALUATION METRICS ARE ACCURACY(ACC), F1-SCORE(F1) AND AUC (ABBREVIATIONS: RW= RECTANGULAR WINDOW, BP= BANDPASS FILTER, MA= MOVING AVERAGE FILTER, KM= KALMAN FILTER, HW= HANNING WINDOW)

Proposed Preprocessing	Metric	Algorithm					
		AB	kNN	LDA	RF	DT	SVM
Without Denoising+ RW	ACC (%)	87.80	85.90	86.86	83.33	78.49	92.69
	F1 (%)	68.91	85.90	79.82	66.88	63.60	90.12
	AUC (%)	82.21	89.80	86.16	78.66	76.78	91.37
BP+MA	ACC (%)	89.99	88.51	92.57	84.57	90.64	92.89
	F1 (%)	80.23	88.51	86.07	78.92	81.60	90.42
	AUC (%)	88.06	93.04	89.06	83.56	87.92	91.56
BP+MA+KM	ACC (%)	91.64	76.08	95.01	85.82	90.90	95.07
	F1 (%)	86.99	76.08	88.46	86.99	80.29	89.68
	AUC (%)	92.01	83.22	91.86	84.96	84.89	92.89
BP+MA+KM +HW	ACC (%)	89.52	81.35	95.04	91.83	93.41	**95.55**
	F1 (%)	85.08	81.35	89.60	82.84	88.27	**91.42**
	AUC (%)	87.92	86.89	92.42	89.71	92.77	**93.25**

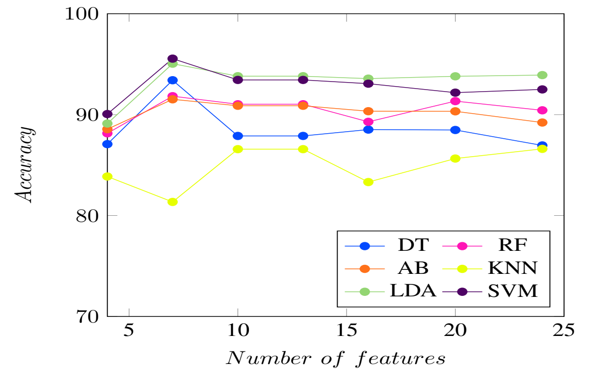

Fig. 4. Features impact on stress detection accuracy for different algorithms

This table summarizes the performance metrics achieved by each algorithm on the stress detection task. In this table we only represent the results of same algorithm in both of our work and P. Schmit *et al.* [15] work, and we did not represent results of SVM in the table. The results of the LDA show that the classifier used is not the only factor for achieving such a good performance. LDA achieved a high accuracy of 95.05% and F1-Score of 89.60%, which is 9.19 % improvement in accuracy and 6.52% improvement in F1-Score in comparison with P. Schmit *et al.* work as the highest results among other papers in Table III. Given that they are both using the same classifier and we use a subset of features that are used in Schmit *et al.* work, this improvement in results is attributed to the improvements thanks to the proposed denoising filters and window segmentation, alongside selecting the best feature set than their work.

V. CONCLUSION

Using biological signals to detect stress is an important step in research on the management of stress and prevention of issues caused by unhealthy stress. We introduced a novel method that utilizes the readily available and non-intrusive PPG sensor embedded in smartwatches. Using advanced denoising techniques, segmentation methods, and extraction of key features, our proposed approach achieves an impressive 95.55% accuracy and 91.42% F1-Score in stress detection using the SVM algorithm, which is the highest in the literature. Avoiding complexity in the process and high accuracy renders this method a suitable candidate for implementation in smartphones, enabling real-time stress detection for a broader applications. In future studies, we plan to explore the selection of the most suitable machine learning algorithms by implementing model selections that optimize hyperparameters and features, thereby enhancing all evaluation metrics.

TABLE III
COMPARING OUR PROPOSED METHOD AGAINST SCHMIT *et al.* [15]

Method	Metric	Algorithm				
		AB	kNN	LDA	RF	DT
Schmit *et al.* [15]	ACC (%)	84.10	82.06	85.83	84.18	81.39
	F1 (%)	81.23	78.90	83.08	81.35	78.27
Proposed Method	ACC (%)	89.52	81.35	**95.04**	91.83	82.84
	F1 (%)	85.08	81.35	**89.60**	78.92	88.27

References

[1] G. Fink, "Stress, definitions, mechanisms, and effects outlined: Lessons from anxiety," in *Stress: Concepts, cognition, emotion, and behavior*. Elsevier, 2016, pp. 3–11.

[2] A. K. Randall and G. Bodenmann, "Stress and its associations with relationship satisfaction," *Current opinion in psychology*, vol. 13, pp. 96–106, 2017.

[3] C. Hammen, E. Y. Kim, N. K. Eberhart, and P. A. Brennan, "Chronic and acute stress and the prediction of major depression in women," *Depression and anxiety*, vol. 26, no. 8, pp. 718–723, 2009.

[4] M. R. Salleh, "Life event, stress and illness," *The Malaysian journal of medical sciences: MJMS*, vol. 15, no. 4, p. 9, 2008.

[5] H. Thapliyal, V. Khalus, and C. Labrado, "Stress detection and management: A survey of wearable smart health devices," *IEEE Consumer Electronics Magazine*, vol. 6, no. 4, pp. 64–69, 2017.

[6] A. Anzanpour, I. Azimi, M. Götzinger, A. M. Rahmani, N. TaheriNejad, P. Liljeberg, A. Jantsch, and N. Dutt, "Self-awareness in remote health monitoring systems using wearable electronics," in *Proceedings of Design and Test Europe Conference (DATE)*, Lausanne, Switzerland, March 2017.

[7] N. TaheriNejad and D. Pollreisz, "Assessment of physiological signals during happiness, sadness, pain or anger," in *6th MobiHealth*, 2016.

[8] P. Perego, A. M. Rahmani, and N. Taherinejad, *Wireless Mobile Communication and Healthcare*, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Springer International Publishing, 2018.

[9] N. TaheriNejad, "Wearable medical devices: Challenges and self-aware solutions," in *IEEE Life Sciences*, vol. 2, April 2019, pp. 5–6. [Online]. Available: https://lifesciences.ieee.org/lifesciences-newsletter/2019/april-2019/wearable-medical-devices-challenges-and-self-aware-solutions/

[10] N. TaheriNejad and Y. Rahmati, "Blood pressure estimation using a single ppg signal," in *2nd EAI International Conference on Wearables in Healthcare (Healthwear 2020)*, Dec 2020, pp. 1–9.

[11] D. Pollreisz and N. TaheriNejad, "Efficient respiratory rate extraction on smartwatch," in *42nd Annual International Conferences of the IEEE Engineering in Medicine and Biology Society in conjunction with the 43rd Annual Conference of the Canadian Medical and Biological Engineering Society EMBC*, Montreal, Canada, July 2020, pp. 1–4.

[12] R. A. Khan and A.-S. K. Pathan, "The state-of-the-art wireless body area sensor networks: A survey," *International Journal of Distributed Sensor Networks*, vol. 14, no. 4, p. 1550147718768994, 2018.

[13] N. TaheriNejad, "Functional or fictional," *IEEE Life Sciences*, vol. Special Issue on: Consumer Technology meets Health Care, pp. 13–15, February 2020.

[14] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 440–460, 2019.

[15] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing wesad, a multimodal dataset for wearable stress and affect detection," in *Proceedings of the 20th ACM international conference on multimodal interaction*, 2018, pp. 400–408.

[16] M. Götzinger, N. Taherinejad, A. M. Rahmani, P. Liljeberg, A. Jantsch, and H. Tenhunen, "Enhancing the early warning score system using data confidence," in *International Conference on Wireless Mobile Communication and Healthcare*. Springer, 2017, pp. 91–99.

[17] D. Sopic, A. Aminifar, A. Aminifar, and D. Atienza, "Real-time event-driven classification technique for early detection and prevention of myocardial infarction on wearable systems," *IEEE transactions on biomedical circuits and systems*, vol. 12, no. 5, pp. 982–992, 2018.

[18] F. Forooghifar, A. Aminifar, L. Cammoun, I. Wisniewski, C. Ciumas, P. Ryvlin, and D. Atienza, "A self-aware epilepsy monitoring system for real-time epileptic seizure detection," *Mobile Networks and Applications*, pp. 1–14, 2019.

[19] M. Götzinger, A. Anzanpour, I. Azimi, N. TaheriNejad, A. Jantsch, A. Rahmani, and P. Liljeberg, "Confidence-enhanced early warning score based on fuzzy logic," *ACM/Springer Mobile Networks and Applications*, pp. 1–18, August 2019.

[20] P. Perego, N. Taherinejad, and M. Caon, *Wearables in Healthcare*, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering book series (LNICST, volume 376). Springer International Publishing, 2021.

[21] A. Aminifar, M. Shokri, F. Rabbi, V. K. I. Pun, and Y. Lamo, "Extremely randomized trees with privacy preservation for distributed structured health data," *IEEE Access*, vol. 10, pp. 6010–6027, 2022.

[22] D. Freismuth and N. TaheriNejad, "On the treatment and diagnosis of attention deficit hyperactivity disorder with eeg assistance," *Electronics Journal*, pp. 1–18, 2022.

[23] D. Pollreisz and N. TaheriNejad, "A simple algorithm for emotion recognition, using physiological signals of a smart watch," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, July 2017, pp. 2353–2356.

[24] V. Montesinos, F. Dell'Agnola, A. Arza, A. Aminifar, and D. Atienza, "Multi-modal acute stress recognition using off-the-shelf wearable devices," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 2196–2201.

[25] A. Aminifar, F. Rabbi, V. K. I. Pun, and Y. Lamo, "Monitoring motor activity data for detecting patients' depression using data augmentation and privacy-preserving distributed learning," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2021, pp. 2163–2169.

[26] A. Jahanjoo, M. Naderan, and M. J. Rashti, "Detection and multi-class classification of falling in elderly people by deep belief network algorithms," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, pp. 4145–4165, 2020.

[27] A. Jahanjoo, M. N. Tahan, and M. J. Rashti, "Accurate fall detection using 3-axis accelerometer sensor and mlf algorithm," in *2017 3rd International Conference on Pattern Recognition and Image Analysis (IPRIA)*, 2017, pp. 90–95.

[28] D. Das, S. Datta, T. Bhattacharjee, A. D. Choudhury, and A. Pal, "Eliminating individual bias to improve stress detection from multimodal physiological data," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 5753–5758.

[29] Y. Hasanpoor, K. Motaman, B. Tarvirdizadeh, K. Alipour, and M. Ghamari, "Stress detection using ppg signal and combined deep cnn-mlp network," in *2022 29th National and 7th International Iranian Conference on Biomedical Engineering (ICBME)*, 2022, pp. 223–228.

[30] M. Benchekroun, B. Chevallier, H. Beaouiss, D. Istrate, V. Zalc, M. Khalil, and D. Lenne, "Comparison of stress detection through ecg and ppg signals using a random forest-based algorithm," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2022, pp. 3150–3153.

[31] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Frontiers in public health*, p. 258, 2017.

[32] R. E. Kalman, "A new approach to linear filtering and prediction problems," 1960.

[33] A. M. Kabe and B. H. Sako, "Chapter 5 - analysis of continuous and discrete time signals," in *Structural Dynamics Fundamentals and Advanced Applications*, A. M. Kabe and B. H. Sako, Eds. Academic Press, 2020, pp. 271–427. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128216156000058

[34] J. Obuchowski, A. Wyłomańska, and R. Zimroz, "The local maxima method for enhancement of time–frequency map and its application to local damage detection in rotating machines," *Mechanical Systems and Signal Processing*, vol. 46, no. 2, pp. 389–405, 2014.

[35] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[36] C. J. Van Rijsbergen, "Foundation of evaluation," *Journal of documentation*, vol. 30, no. 4, pp. 365–373, 1974.

[37] T. Fawcett, "An introduction to roc analysis," *Pattern recognition letters*, vol. 27, no. 8, pp. 861–874, 2006.