A Novel Machine-Learning-Based Noise Detection Method for Photoplethysmography Signals

Soheil Khooyooz[1], Anice Jahanjoo[2], Amin Aminifar[1], Nima TaheriNejad[1,2]
[1]Heidelberg University, Germany   [2]Technische Universität Wien (TU Wien), Austria

Abstract—Wearable devices are widespread for continuous health monitoring; capturing various physiological parameters for remote health monitoring and early detection of health issues. These devices are susceptible to interference such as Motion Artifacts (MA) and Baseline Wanders (BW). Mitigating potential false alarms due to those artifacts is an important challenge in wearable healthcare. To tackle this challenge, it is crucial to first identify noise in the signals recorded by wearable systems. Most of the conventional methods rely on reference data like accelerometer data to detect noise in Photoplethysmogram (PPG) signals. This study proposes a Machine Learning (ML)-based approach to distinguish between clean and corrupted segments in PPG signals without relying on other sensors' data. Binary and three-class classification on clean, MA-, and BW-corrupted signals produce promising F1-scores from 89.3% to 99.4%.

Index Terms—earable devices; health monitoring; photo-plethysmography; noise detection; motion artifact; baseline wanderearable devices; health monitoring; photoplethysmography; noise detection; motion artifact; baseline wanderW

I. INTRODUCTION

Nowadays, wearable devices have become very common for continuous health monitoring. Using wearables, we can record various physiological parameters from the body, which can be utilized for remote health monitoring and early detection of health problems [1]–[5]. While wearable devices have facilitated and improved self-care and healthcare quality, they are exposed to several different types of noise, including motion and muscle artifact, Baseline Wanders (BW), and power-line interference [6].

When a wearable device exhibits a high rate of false positives, patients may not adequately notice the system's alarms. Conversely, in specific applications such as detecting atrial fibrillations or sleep apneas, increased false negative rates could lead to significant repercussions. Thus, decreasing the occurrence of false positives emerges as a notable challenge within the wearable healthcare domain [7].

To address this challenge, it is important to initially identify the presence of noise in the signals recorded by the wearable system. In [8], for atrial fibrillation detection, the authors employed a variable frequency complex demodulation (VFCDM) approach to identify noise artifacts in the signals captured from fingertip videos of smartphones. In [5], to enhance atrial fibrillation detection, the authors used an autoencoder for Motion Artifact (MA) detection and removal from Photoplethysmogram (PPG) signals recorded by smartwatches. In

[9], the authors introduced a fuzzy-logic-based approach to enhance the early warning score (EWS) systems, essential for monitoring the general health status of patients.

In [10], the authors introduced a probabilistic neural network method to address motion and noise artifacts (MNAs) affecting PPG signals obtained from smartphone cameras during health measurements. In [11], the authors proposed a statistical approach involving the calculation of kurtosis and Shannon entropy to identify motion and noise artifacts in PPG signals. In [12], the authors conducted comprehensive research on various approaches aimed at enhancing the reliability of physiological parameter measurements, with a particular focus on PPG signals obtained from wearable devices. The study involved a thorough review of state-of-the-art algorithms specifically designed for detecting MA in PPG signals. Lastly, in [13], the author introduced several time and time-frequency domain approaches for detecting noise and artifacts, as well as signal reconstruction algorithms in wearable devices based on Electrocardiography (ECG) and PPG.

One issue with most of the previous methods is dependence on various sensors, such as accelerometer sensors, for noise detection in PPG signals. This dependency increases design complexity and computational requirements. Conversely, the Machine Learning (ML)-based approaches offer a more precise detection of various types of noise sources [14], [15].

In this paper, we introduce a ML-based approach for detecting clean, MA-, and BW-corrupted segments in PPG signals without relying on other sensors' data. Initially, clean and noisy signal segments are identified using an annotation algorithm applied to the dataset. Subsequently, morphological and statistical features of these signals are extracted. Finally, various ML algorithms are trained using these features as predictive models.

The rest of this paper is organized as follows. In Section II, we introduce the dataset used for training and evaluating our predictive model. Section III elaborates on our methodology. In Section IV, the evaluation metrics and the experimental results are discussed. Finally, in Section V, we conclude our findings for this paper.

II. DATASET

In this paper, we use an open-access dataset on emotion, cognition, sleep, and multi-model physiological signals (EC-SMP) [16]. The dataset consists of multi-modal physiological signals, including Electroencephalography (EEG), ECG, PPG,
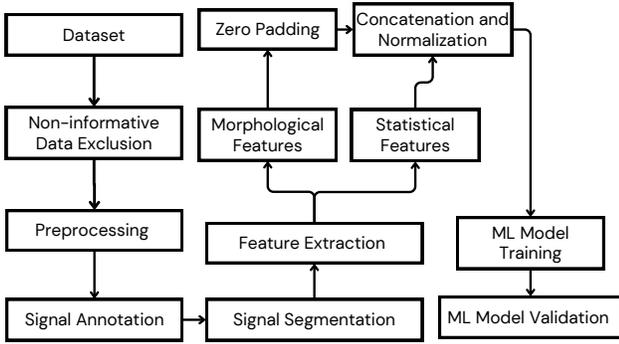
Fig. 1: Block diagram of the proposed method

Electrodermal Activity (EDA), temperature, and accelerometer data, recorded from 89 healthy college students during resting state, different emotion induction and recovery, and cognitive function assessment tasks. We utilize the data collected using the Empatica E4 smart-watch, which includes PPG, EDA, accelerometer, and temperature signals. Among these, we use PPG and accelerometer signals for our research, in which the accelerometer data is only used for the data preparation phase. The E4 files contain data from only 67 subjects (24 males and 43 females, age: 23.82 ± 1.93 years), each contributing an average of nearly one and a half hours. The PPG and the accelerometer signals have a sampling rate of 64 and 32 samples per second, respectively.

III. METHODOLOGY

In this section, we elaborate on our methods and algorithms employed for noise recognition in PPG signals. The PPG signals obtained from the ECSMP dataset are raw signals containing offset and MA. Due to the absence of references for the offset parts of the signals to be used as BW, we filter the signal with a highpass filter to remove the offsets. We then utilize an annotation algorithm to identify clean and MA-corrupted segments of the signals, using the accelerometer signal as our reference. Subsequently, we extract various features from the segmented signals and train our predictive models based on these features. It is worth to mention that in the train and test separation stage, we split the dataset into 80% for training and 20% for testing. To ensure a fair classification, we use data from a specific subject exclusively for either testing or training. The predictive models are trained using the training data, and the models are subsequently evaluated using the test data. A block diagram illustrating this approach is presented in Fig. 1.

A. Preprocessing

In the preprocessing phase, we start by removing non-informative segments at the end of each subject's recorded data, captured when the devices are no longer worn. Subsequently, to eliminate the unknown offset, we apply a highpass filter to the data, using a 0.5 Hz 4th order IIR filter [12], [17].

B. Annotation algorithm

To prepare the data for the feature extraction step, we extract clean and MA-corrupted segments from each subject's signals using the following algorithm:

1) Calculation of the noise reference: To encompass every potential motion that may corrupt the PPG signal, we utilize a reference signal derived from accelerometer data, calculated as follows:

$$a = \sqrt{a_x^2 + a_y^2 + a_z^2} \qquad (1)$$

in which a_x, a_y, and a_z are the accelerations recorded by the accelerometer sensor in the x, y, and z directions.

2) Noise reference energy calculation: Energy serves as a valuable metric for detecting the level of motion in the accelerometer signal. Hence, we compute the energy of the samples in the signal a. Samples surpassing a specific energy threshold are identified as motion, and their counterparts in the PPG signal are annotated as motion samples. To amplify the corrupted samples and diminish the uncorrupted ones, we first calculate the derivative of the signal a, then calculate its energy [18]:

$$re[n] = (a[n] - a[n-1])^2 \qquad (2)$$

in which re and n are the energy of the reference signal and the number of the sample in the accelerometer signal, respectively.

3) Binary signal calculation: After the energy of the reference is calculated, we apply a threshold on the signal to find clean and corrupted samples:

$$bs[n] = \begin{cases} 1 & \text{if } re[n] > th \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

in which th is calculated as follows:

$$th = 0.5 \times \frac{1}{67} \sum_{i=1}^{67} a_i \qquad (4)$$

Which describes the average of the accelerometer signals across all subjects multiplied by 0.5.

4) Final segmentation: Based on the calculated binary signal in step 3, for each subject, we extract clean and noisy segments using a non-overlapping window of size 1280 samples (20 seconds). This number of samples is sufficient for assessing the PPG signal to determine if it is clean or corrupted with noise. Consequently, if, for 1280 successive samples, the values of the binary signal remain 0/1, we consider the corresponding window in the PPG signal as a clean/noisy segment; otherwise, we ignore the window (because it includes both clean and noisy signals). After segmentation, we balance the number of segments based on the minority class to conduct a fair classification. The results of this algorithm are depicted in Fig. 2.

C. BW synthesis

To investigate the performance of our model, we also synthesize random BW using a combination of sine and cosine waves. BW is primarily caused by respiration at frequencies ranging between 0.15 to 0.3 Hz (0.5 Hz) [19], [20]. Therefore,

(a) annotated accelerometer signal

(b) clean segment
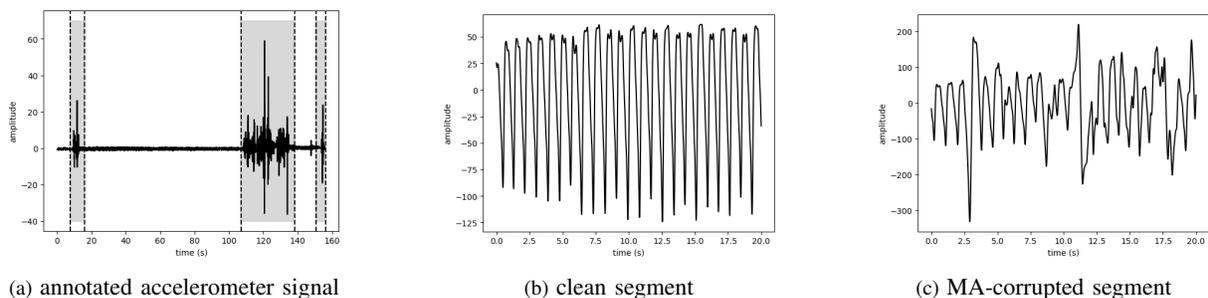
(c) MA-corrupted segment

Fig. 2: Annotation algorithm's results: (a) MA-corrupted parts are shown in the grayed-out sections in the accelerometer signal (b) extracted clean PPG signal (c) extracted MA-corrupted signal
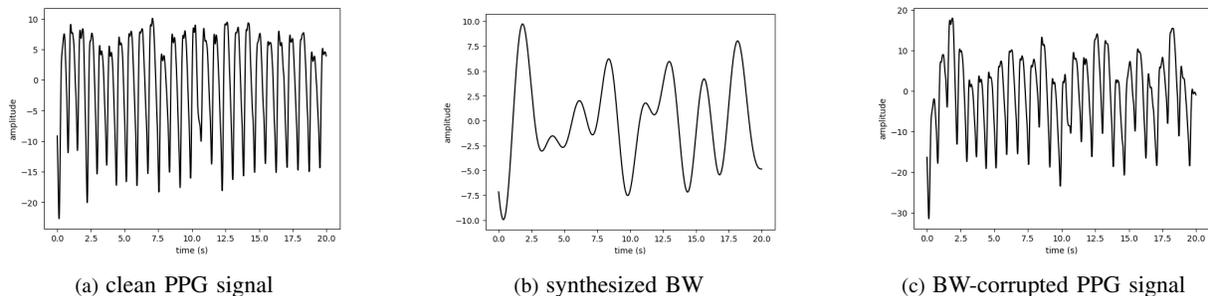

(a) clean PPG signal

(b) synthesized BW

(c) BW-corrupted PPG signal

Fig. 3: Singal types: (a) clean PPG signal (b) synthesized BW signal (c) BW-corrupted PPG signal

(a) clean PPG signal

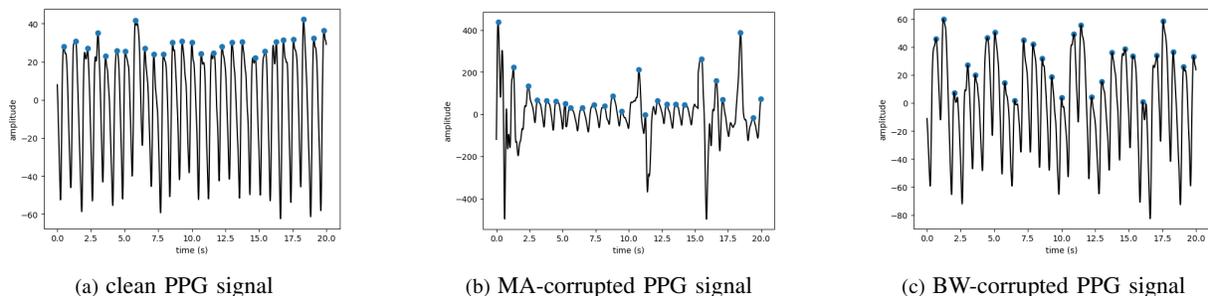(b) MA-corrupted PPG signal

(c) BW-corrupted PPG signal

Fig. 4: Peak detection results using vital-sqi peak detection algorithm for (a) clean, (b) MA-corrupted PPG signal, and (c) BW-corrupted PPG signal

we select the frequency arguments of the sinusoidal waves and their amplitudes to be within the range of 0.15 to 0.3 Hz (0.5 Hz) and 0.1 to 0.25 (0.5) times the maximum amplitude of each window of data, respectively. Subsequently, we add these random signals to the clean data and consider them as BW-corrupted data. A sample of these synthesized artifacts and their corresponding noisy segments are depicted in Fig. 3.

D. Feature extraction

In this research, we conduct noise detection based on morphological and statistical features.

1) Morphological features: Morphological features describe the shape, pattern, and specific characteristics of signals. Since noise can alter the characteristics of the PPG signal in terms of shape and pattern, we calculate the locations of peaks in the signal using the vital-sqi open-access Python toolbox [21], as illustrated in Fig. 4. The distances between the arguments of these peaks (time intervals between peaks)

are then considered as morphological features. For different classes, the distance pattern between detected peaks varies. For example, in MA-corrupted signals, some peaks may go undetected, resulting in a distinct distance pattern compared to other classes. To ensure uniformity in feature vectors, we follow these steps: 1. Sort the features in each feature vector in descending order. 2. Calculate the mean (μ) and standard deviation (σ) of the vectors. 3. Set the length of the feature vectors to $l = \lfloor \mu \rfloor + \lfloor \sigma \rfloor$. 4. Zero-pad feature vectors smaller than l, and ignore the last samples in feature vectors larger than l [7].

2) Statistical features: Statistical features encompass the distribution and variability of signals. To capture this information, we calculate the mean, variance, kurtosis, skewness, energy, entropy, and maximum auto-correlation of the signal, along with the mean, variance, and maximum of the signal's histogram [22], [23].

E. Classification

In this study, we conduct binary and three-class classifications as follows: a) Clean vs. MA-corrupted data, b) Clean vs. BW-corrupted data, c) MA- vs. BW-corrupted data, d) Clean vs. MA- vs. BW-corrupted data.

To achieve this, we employ Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel [24], Random Forest (RF) [25], XGBoost (XGB) [26], Naive Bayes (NB) [27], Extremely Randomized Tree (ERT) [28], and Decision Trees (DT) [29] classifiers to train our predictive model. We implement these classifiers using scikit-learn [30] with default parameters.

IV. EVALUATION

A. Evaluation Metrics

For performance evaluation of our ML models, we use the following metrics:

$$F1\text{-}score = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{5}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}, \tag{6}$$

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{7}$$

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{8}$$

in which TP, TN, FP, and FN are true positives, true negatives, false positives, and false negatives, respectively. Before training and validating our models, we ensure data balance across all three classes. Hence, all metrics are computed with equal weight for the various classes.

B. Experimental Results

We conduct classifications on the dataset using a 5-fold cross-validation approach in which each time 80% of subjects are used for train and 20% of them are used for test. We conduct both binary and three-class classifications. The experiment is conducted for both of the two different ranges for BW as reported in the literature. The results for F1-score, Accuracy, Precision, and Recall are provided in Table I and II.

As shown in Tables I and II, various machine learning models are employed for classification. For instance, in the "Clean vs. MA" comparison, the ERT model outperforms other methods, achieving an F1-score of 97.0% ± 0.6% and an accuracy of 96.9% ± 0.6%. Similar performance trends are observed in the "Clean vs. BW," "MA vs. BW," and "Clean vs. MA vs. BW" comparisons, for the XGB and ERT algorithms across different BW ranges, confirming the noise robustness of these algorithms compared to others [31]. Additionally, it is evident that wider BW ranges lead to improved results, as shown in Fig. 5.

V. CONCLUSION

Wearable devices are susceptible to various forms of noise, which can adversely impact their functionality and lead to inaccurate detection. To address this challenge, we propose a noise detection model that employs diverse ML algorithms

(a) Clean vs. MA (b) Clean vs. BW
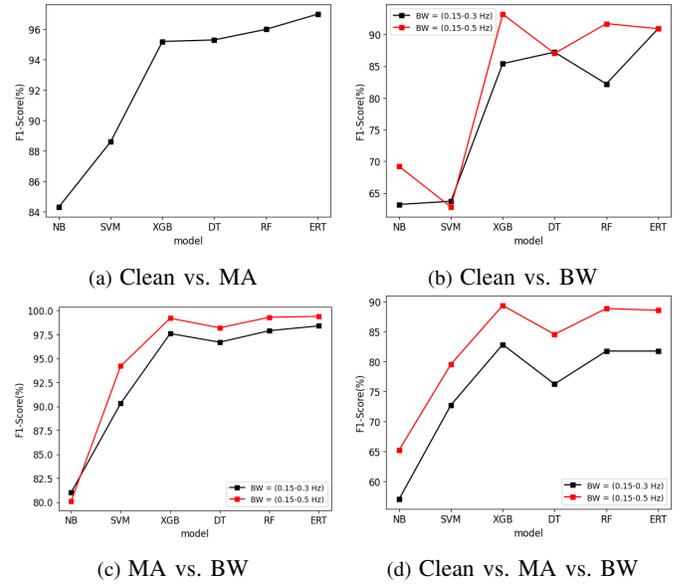
(c) MA vs. BW (d) Clean vs. MA vs. BW

Fig. 5: Results comparison for different ML models and BW ranges

TABLE I: Classification results for various comparisons and BW, with a frequency range of 0.15 to 0.3 Hz and amplitude ranging from 0.1 to 0.25 times the maximum amplitude of PPG segments

Comparison	Metrics			
	Acc.(%)	Rec.(%)	Prec.(%)	F1 (%)
Clean vs. MA				
Ml Alg.				
NB	84.1 ± 1.5	83.4 ± 2.7	85.4 ± 0.8	84.3 ± 1.2
SVM	88.2 ± 1.3	85.8 ± 2.5	91.7 ± 0.8	88.6 ± 1.1
XGB	95.0 ± 1.0	91.3 ± 1.6	99.5 ± 0.2	95.2 ± 0.9
DT	95.1 ± 0.7	91.5 ± 1.2	99.5 ± 0.2	95.3 ± 0.7
RF	95.6 ± 0.9	92.4 ± 1.5	99.7 ± 0.1	96.0 ± 0.8
ERT	96.9 ± 0.6	94.5 ± 1.1	99.6 ± 1.8	97.0 ± 0.6
Clean vs. BW				
NB	65.0 ± 1.0	66.8 ± 2.0	60.0 ± 3.2	63.2 ± 1.3
SVM	53.8 ± 3.2	52.5 ± 2.1	81.0 ± 3.2	63.7 ± 2.5
XGB	85.2 ± 1.3	84.3 ± 1.8	86.5 ± 2.8	85.4 ± 1.4
DT	87.2 ± 1.0	87.4 ± 1.6	87.0 ± 2.4	87.2 ± 1.1
RF	82.0 ± 1.2	81.2 ± 2.2	83.4 ± 4.1	82.2 ± 1.5
ERT	90.9 ± 1.4	91.1 ± 1.8	91.0 ± 2.8	90.9 ± 1.5
MA vs. BW				
NB	80.7 ± 1.9	79.9 ± 3.1	82.1 ± 0.8	81.0 ± 1.5
SVM	90.4 ± 1.1	91.2 ± 1.8	89.5 ± 0.5	90.3 ± 1.1
XGB	97.6 ± 0.6	95.6 ± 1.0	99.7 ± 0.2	97.6 ± 0.5
DT	96.6 ± 0.5	94.0 ± 0.9	99.5 ± 0.2	96.7 ± 0.5
RF	97.8 ± 0.5	96.1 ± 1.0	99.7 ± 0.2	97.9 ± 0.5
ERT	98.3 ± 0.5	97.2 ± 0.9	99.6 ± 0.2	98.4 ± 0.5
Clean vs. MA vs. BW				
NB	59.4 ± 13.1	59.5 ± 6.7	59.4 ± 22.6	57.0 ± 14.9
SVM	72.8 ± 0.8	72.9 ± 5.0	72.8 ± 7.3	72.7 ± 5.5
XGB	82.9 ± 0.5	82.9 ± 4.4	82.9 ± 7.7	82.8 ± 5.7
DT	76.3 ± 0.1	76.1 ± 9.4	76.3 ± 11.5	76.2 ± 10.3
RF	81.8 ± 0.5	81.8 ± 6.2	81.8 ± 9.0	81.7 ± 7.2
ERT	81.7 ± 0.6	81.8 ± 7.8	81.7 ± 8.7	81.7 ± 7.9

to identify segments of PPG signals corrupted by noise. In contrast to previous works, we exclusively rely on PPG signals for noise detection, making the procedure more straightfor-

TABLE II: Classification results for various comparisons and BW, with a frequency range of 0.15 to 0.5 Hz and amplitude ranging from 0.1 to 0.5 times the maximum amplitude of PPG segments

Comparison	Metrics			
	Acc.(%)	Rec.(%)	Prec.(%)	F1 (%)
Clean vs. MA				
Ml Alg.				
NB	84.1 ± 1.5	83.4 ± 2.7	85.4 ± 0.8	84.3 ± 1.2
SVM	88.2 ± 1.3	85.8 ± 2.5	91.7 ± 0.8	88.6 ± 1.1
XGB	95.0 ± 1.0	91.3 ± 1.6	99.5 ± 0.2	95.2 ± 0.9
DT	95.1 ± 0.7	91.5 ± 1.2	99.5 ± 0.2	95.3 ± 0.7
RF	95.6 ± 0.9	92.4 ± 1.5	99.7 ± 0.1	96.0 ± 0.8
ERT	96.9 ± 0.6	94.5 ± 1.1	99.6 ± 1.8	97.0 ± 0.6
Clean vs. BW				
NB	72.7 ± 1.2	79.6 ± 3.9	68.8 ± 3.4	69.2 ± 1.2
SVM	48.6 ± 2.0	49.2 ± 1.1	48.7 ± 2.5	62.8 ± 1.5
XGB	93.2 ± 1.2	93.0 ± 1.8	93.2 ± 1.6	93.2 ± 1.2
DT	87.0 ± 1.0	87.1 ± 1.7	87.0 ± 2.0	87.0 ± 1.1
RF	91.7 ± 1.2	91.7 ± 1.8	91.6 ± 1.3	91.7 ± 1.2
ERT	91.0 ± 1.1	91.6 ± 1.8	90.2 ± 1.8	90.9 ± 1.1
MA vs. BW				
NB	79.7 ± 0.3	78.8 ± 0.4	81.6 ± 0.5	80.1 ± 0.3
SVM	94.3 ± 0.2	96.8 ± 0.2	91.7 ± 0.1	94.2 ± 0.2
XGB	99.1 ± 0.1	98.4 ± 0.3	99.9 ± 0.1	99.2 ± 0.1
DT	98.2 ± 0.3	96.8 ± 0.5	99.7 ± 0.2	98.2 ± 0.2
RF	99.3 ± 0.1	98.8 ± 0.2	99.8 ± 0.1	99.3 ± 0.1
ERT	99.4 ± 0.2	99.0 ± 0.3	99.8 ± 0.1	99.4 ± 0.2
Clean vs. MA vs. BW				
NB	65.7 ± 1.5	68.1 ± 8.1	65.2 ± 13.8	65.2 ± 6.9
SVM	79.5 ± 0.9	79.6 ± 4.4	79.5 ± 4.0	79.5 ± 0.9
XGB	89.3 ± 0.7	89.3 ± 6.9	89.3 ± 3.6	89.3 ± 3.6
DT	84.5 ± 0.4	84.5 ± 5.0	84.5 ± 6.6	84.5 ± 5.5
RF	88.8 ± 0.7	88.8 ± 3.2	88.7 ± 5.3	88.8 ± 3.8
ERT	88.5 ± 0.7	88.5 ± 4.5	88.5 ± 4.1	88.5 ± 6.6

ward. Additionally, our approach introduces potential ideas for simplifying the hardware design in future wearable devices. By extracting both morphological and statistical features from PPG signals, our approach trains ML models to distinguish between clean and corrupted segments. The outcomes of our method are promising, with F1-scores of 97.0%, 93.2%, 99.4%, and 89.3% achieved for the classifications of clean vs. MA-corrupted signals, clean vs. BW-corrupted signals, MA- vs. BW-corrupted signals, and clean vs. MA- vs. BW-corrupted signals, respectively.

REFERENCES

[1] X. Erickson *et al.* Personalized seizure detection using spiking neural networks. In *2023 IEEE International Conference on Omni-layer Intelligent Systems (COINS)*, pp. 1–6. IEEE, 2023.

[2] 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*, pp. 1–4, Montreal, Canada, July 2020.

[3] A. Aminifar *et al.* 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)*, pp. 2163–2169. IEEE, 2021.

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

[5] F. Mohagheghian *et al.* Atrial fibrillation detection on reconstructed photoplethysmography signals collected from a smartwatch using a denoising autoencoder. *Expert Systems With Applications*, 237:121611, 2024.

[6] G. B. Moody *et al.* A noise stress test for arrhythmia detectors. *Computers in cardiology*, 11(3):381–384, 1984.

[7] A. Aminifar *et al.* Recognoise: Machine-learning-based recognition of noisy segments in electrocardiogram signals. In *2024 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2024.

[8] S. K. Bashar *et al.* Developing a novel noise artifact detection algorithm for smartphone ppg signals: Preliminary results. In *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pp. 79–82. IEEE, 2018.

[9] M. Götzinger *et al.* Confidence-enhanced early warning score based on fuzzy logic. *ACM/Springer Mobile Networks and Applications*, pp. 1–18, August 2019.

[10] F. Tabei *et al.* A novel personalized motion and noise artifact (mna) detection method for smartphone photoplethysmograph (ppg) signals. *IEEE Access*, 6:60498–60512, 2018.

[11] N. Selvaraj *et al.* Statistical approach for the detection of motion/noise artifacts in photoplethysmogram. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4972–4975. IEEE, 2011.

[12] D. Pollreisz and N. TaheriNejad. Detection and removal of motion artifacts in ppg signals. *Mobile Networks and Applications*, pp. 1–11, 2019.

[13] S. M. A. Salehizadeh. Motion and noise artifact detection and vital signal reconstruction in ecg/ppg based wearable devices. 2015.

[14] A. Olyanasab and M. Annabestani. Leveraging machine learning for personalized wearable biomedical devices: A review. *Journal of Personalized Medicine*, 14(2):203, 2024.

[15] H. C. Ates *et al.* End-to-end design of wearable sensors. *Nature Reviews Materials*, 7(11):887–907, 2022.

[16] Z. Gao *et al.* Ecsmp: A dataset on emotion, cognition, sleep, and multi-model physiological signals. *Data in Brief*, 39:107660, 2021.

[17] A. Jahanjoo *et al.* High-accuracy stress detection using wrist-worn ppg sensors. In *2024 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2024.

[18] S. Khooyooz *et al.* Fabrication of a low-cost real-time mobile ecg system for health monitoring. *arXiv preprint arXiv:2302.06272*, 2023.

[19] F. A. Orakzai *et al.* Baseline wandering removal from human electrocardiogram signal using projection pursuit gradient ascent algorithm. *International Journal of Electrical & Computer Sciences IJECS/IJENS*, 9(09):351–354, 2009.

[20] A. E. Awodeyi *et al.* Median filter approach for removal of baseline wander in photoplethysmography signals. In *2013 European Modelling Symposium*, pp. 261–264. IEEE, 2013.

[21] V.-K. D. Le *et al.* vital_sqi: A python package for physiological signal quality control. *Frontiers in Physiology*, pp. 2248, 2022.

[22] S. Khooyooz and S. H. Sardouie. Classification of motor and mental imagery eeg signals in bci systems based on signal-to-image conversion. In *2022 29th National and 7th International Iranian Conference on Biomedical Engineering (ICBME)*, pp. 124–128. IEEE, 2022.

[23] R. Krishnan *et al.* Analysis and detection of motion artifact in photoplethysmographic data using higher order statistics. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 613–616. IEEE, 2008.

[24] C. Cortes and V. Vapnik. Support-vector networks. *Machine learning*, 1995.

[25] T. Ho. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*. IEEE, 1995.

[26] T. Chen and C. Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016.

[27] I. Rish et al. An empirical study of the naive bayes classifier. In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, volume 3, pp. 41–46, 2001.

[28] P. Geurts *et al.* Extremely randomized trees. *Machine learning*, 2006.

[29] J. Quinlan. Induction of decision trees. *Machine learning*, 1986.

[30] F. Pedregosa *et al.* Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[31] D. Datta *et al.* Computational intelligence for observation and monitoring: a case study of imbalanced hyperspectral image data classification. *Computational Intelligence and Neuroscience*, 2022, 2022.