

Article

Detection and Characterization of Physical Activity and Psychological Stress from Wristband Data

Mert Sevil ¹, Mudassir Rashid ², Mohammad Reza Askari ², Zacharie Maloney ¹, Iman Hajizadeh ² and Ali Cinar ^{1,2,*}

¹ Department of Biomedical Engineering, Illinois Institute of Technology, Chicago, IL 60616, USA; msevil@hawk.iit.edu (M.S.); zmaloney@hawk.iit.edu (Z.M.)

² Department of Chemical and Biological Engineering, Illinois Institute of Technology, Chicago, IL 60616, USA; mrashid3@iit.edu (M.R.); maskari@hawk.iit.edu (M.R.A.); ihajizad@hawk.iit.edu (I.H.)

* Correspondence: cinar@iit.edu; Tel.: +1-312-567-2042

Received: 3 August 2020; Accepted: 13 October 2020; Published: 4 December 2020



Abstract: Wearable devices continuously measure multiple physiological variables to inform users of health and behavior indicators. The computed health indicators must rely on informative signals obtained by processing the raw physiological variables with powerful noise- and artifacts-filtering algorithms. In this study, we aimed to elucidate the effects of signal processing techniques on the accuracy of detecting and discriminating physical activity (PA) and acute psychological stress (APS) using physiological measurements (blood volume pulse, heart rate, skin temperature, galvanic skin response, and accelerometer) collected from a wristband. Data from 207 experiments involving 24 subjects were used to develop signal processing, feature extraction, and machine learning (ML) algorithms that can detect and discriminate PA and APS when they occur individually or concurrently, classify different types of PA and APS, and estimate energy expenditure (EE). Training data were used to generate feature variables from the physiological variables and develop ML models (naïve Bayes, decision tree, k-nearest neighbor, linear discriminant, ensemble learning, and support vector machine). Results from an independent labeled testing data set demonstrate that PA was detected and classified with an accuracy of 99.3%, and APS was detected and classified with an accuracy of 92.7%, whereas the simultaneous occurrences of both PA and APS were detected and classified with an accuracy of 89.9% (relative to actual class labels), and EE was estimated with a low mean absolute error of 0.02 metabolic equivalent of task (MET). The data filtering and adaptive noise cancellation techniques used to mitigate the effects of noise and artifacts on the classification results increased the detection and discrimination accuracy by 0.7% and 3.0% for PA and APS, respectively, and by 18% for EE estimation. The results demonstrate the physiological measurements from wristband devices are susceptible to noise and artifacts, and elucidate the effects of signal processing and feature extraction on the accuracy of detection, classification, and estimation of PA and APS.

Keywords: signal processing; wristband data; machine learning; physical activity classification; acute psychological stress detection

1. Introduction

Monitoring physical activity (PA) and acute psychological stress (APS) throughout daily life is important in the management of chronic diseases because regular PA can promote cardiovascular health, whereas episodes of APS can increase the risks of adverse cardiovascular events. Wearable device sensors continuously measure multiple physiological variables to enable self-monitoring of health and preventive medicine [1–6]. These signals provide valuable information in real time and act as surrogates for reporting variations in the levels of hormones such as cortisol,

lactate, and epinephrine, which cannot be measured in real-time, noninvasively, and in daily living, to indicate PA and APS [7–12]. Physiological measurements are also useful in automated medical intervention decisions in chronic diseases. For example, in diabetes, PA and APS may affect blood glucose concentrations in opposite directions. Signals from a wearable device would complement the information received from a continuous glucose monitoring device and provide advance information of the presence of PA and/or APS, which will ultimately affect the glucose level, enabling better insulin dosing decisions [6,9,13–16]. The convenience of noninvasive wearable sensors means that the devices can be worn continuously in daily living to monitor the PA of users without hindering or limiting the motions of the users. However, the signals from wearable devices such as wristbands are corrupted by noise and artifacts. They require powerful signal processing algorithms to extract reliable information from noisy data and eliminate the effects of artifacts.

The physiological variables collected from wearable devices have been useful in noninvasive detection of PA and APS [7–11]. Recent developments in signal processing of wearable device biosignals and machine learning (ML) algorithms enabled the integrated analysis of PA and APS by enabling the detection and discrimination of concurrent incidences of PA and APS [17,18]. This was made possible by clinical experiments designed to enrich the training data with various PA and APS inducements, multiple physiological biosignals measured using a single convenient wearable device, and recent advances in signal processing, feature extraction, and ML [17–21]. The recent developments in signal processing are particularly important since noise and artifacts are routinely encountered in real-world data, which can easily mask the differentiating features among the possibly simultaneous incidences of PA and APS [19,22–26]. Despite the significant impact of signal processing on the accuracy of the algorithms, the effects of signal processing on the performance of ML algorithms is often not reported in studies on the detection and discrimination of PA and APS [17,18].

The role of signal processing is even more critical in discriminating between concurrent PA and APS because they may result in similar responses in measurable physiological variables [17,18]. The psychological stressors prompt the activation of various physiological systems that result in the overall stress response with the aim of restoring homeostasis [7,27,28]. The response to APS is typically coordinated by the hypothalamus through the activation of the autonomic nervous system and the pituitary and adrenal glands in the hypothalamus–pituitary–adrenal axis, resulting in the release of stress hormones such as catecholamines and cortisol [27,28]. The stress hormones are difficult to measure continuously in free-living ambulatory conditions [7]. This necessitated research into surrogate biomarkers of APS, including readily measurable physiological biosignals such as heart rate (HR), respiration rate, pupil diameter, skin temperature (ST), electrodermal activity (EDA, also known as galvanic skin response (GSR)), blood volume pulse (BVP), electrocardiogram, and blood pressure [7,11,13,29,30]. The validity of these conveniently measured physiological biosignals for detecting APS has been demonstrated in several studies [7,11,13,19,29–31]. Previous works illustrated that reliably discriminating between different types of APS is possible, including emotional anxiety stress and mental stress [7,11,32,33]. Figure 1 illustrates the structure of use of biosignals for APS detection, PA classification, and for various healthcare and public health research applications.

The proposed work consists of signals processing, feature extraction, data preparation, machine learning algorithm development, and evaluation of results (Figure 2). To achieve this aim, we designed experimental protocols to collect data, and we developed ML algorithms to detect APS in the presence of PA, which is a challenging problem as the readily-measurable physiological biosignals used to detect APS, such as HR and GSR, are also affected by PA [13]. The challenges in discriminating among different types of APS during periods of possibly simultaneous occurrences of PA are not only due to the overlapping responses in the physiological measurements, but also due to the noise and artifacts in the biosignals measured from wearable devices, which necessitates effective data filtering and adaptive noise cancellation (ANC) algorithms to extract and enhance the informative signal from the measurements for use in ML algorithms [7,13,17–19].

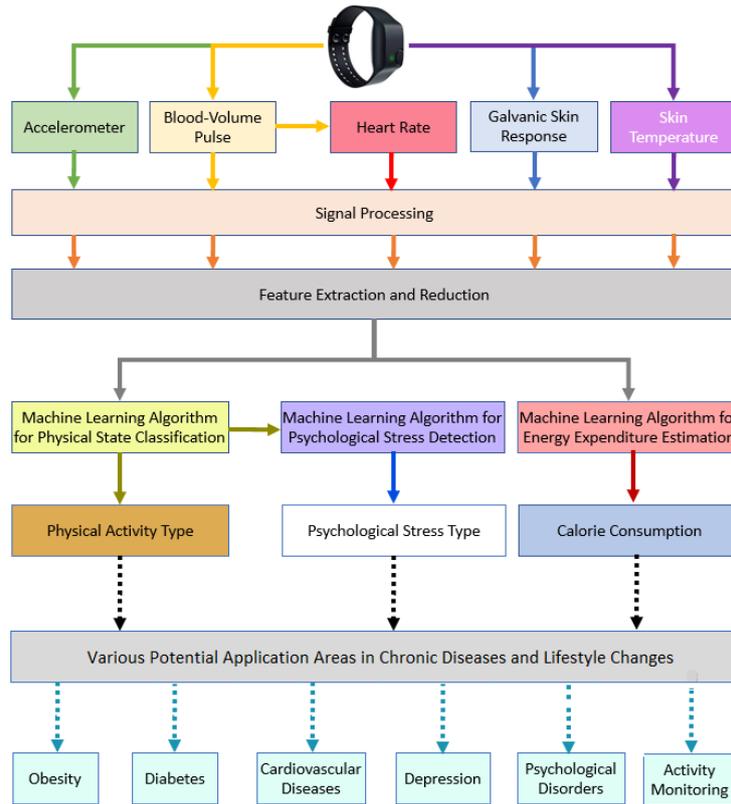


Figure 1. Flowchart of physiological and psychological state estimation using wristband biosignals.

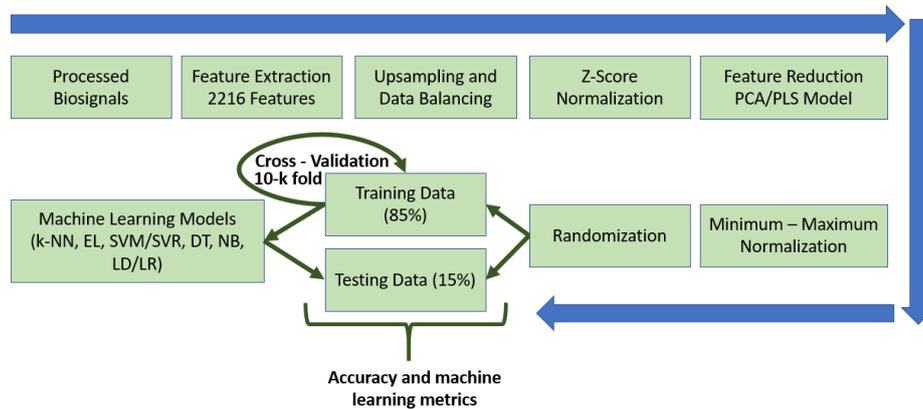


Figure 2. Flowchart of proposed approach.

The practicality of simultaneous PA and APS detection extends beyond routine monitoring of health. The detection and discrimination of simultaneous PA and APS can generate reliable digital biomarkers of the physiological and psychological states of people for use in the treatment of chronic conditions like diabetes and cardiovascular disease [34,35]. Diabetes treatment, in particular, can significantly benefit from more accurate assessments of the PA and APS states of people with Type 1 diabetes (T1D) [9,13–16]. People with T1D must administer exogenous insulin to compensate for the loss of pancreatic insulin production and maintain their blood glucose concentration within the desired safe range [36]. However, insulin requirements vary due to the type and intensity of PA and the possibly concurrent incidence of APS [9,13–16,37]. This creates difficulties for people with T1D in effectively regulating their glucose levels because PA and APS typically have divergent effects on glucose levels [9,13,14,37]. Prolonged low- and moderate-intensity aerobic exercise causes a reduction in glucose concentrations because of the increase in glucose use by the working muscles and

a heightened sensitivity to insulin as muscle cells are more effective using any available insulin to take up glucose during and after PA [9,14,37]. In contrast to PA, APS causes the release of neuroendocrine hormones that stimulate the release of free energy and restricts the uptake of glucose into various tissues, causing the blood glucose concentration to temporarily rise [38]. This can be problematic for regulating glucose levels in people with T1D. Tasks considered routine by healthy individuals, like training or recreational activity versus competitive sporting events, can cause glucose values to drift out of the desired range and increase the risk of developing diabetes-related complications. Some of the routine activities in everyday life, like running due to being late for important events, or undesirable hindrances or disruptions in common tasks can also trigger APS, which may be easily masked by the concurrent PA [13,17,18]. Appropriate signal processing and feature extraction can improve the accuracy of discriminating between physical activity and psychological stress. The necessity of feature extraction in ML is well established [31]. The development of signal processing techniques to handle the noise and artifacts in the measured physiological biosignals from wearable devices is an active research area [19,22–26].

Motivated by the above considerations, in this work, we studied the effects of data filtering and adaptive noise cancellation techniques on the accuracy of detecting and discriminating PA and APS, and quantifying the PA intensity using a variety of ML algorithms and physiological measurements collected from a wristband. We demonstrate that effective signal processing and feature extraction are important to ensure high accuracy for ML algorithms (naïve Bayes, decision tree, k-nearest neighbor, linear discriminant, ensemble learning, and support vector machine) to discriminate among different types of individual or concurrent incidences of PA and APS and quantify the intensity of the PA through estimates of energy expenditure (EE). The PA and APS classification provides users with important information on the individual or concurrent simultaneous occurrences of physiological and psychological stressors that they are experiencing, and EE provides an assessment of the physical exertion. The results show that the proposed signal processing techniques increase the detection and discrimination accuracy for PA and APS and decrease the EE estimation error.

2. Materials and Methods

2.1. Data Collection

Twenty-four subjects participated in 207 different experiments, which consisted of three different physical states (PS) (sedentary state (SS), and two different exercises: treadmill run (TR) and stationary bike (SB)) and three psychological states (non-stressful (NS) and two different APS inducements: exciting-anxiety stress (EAS) and mental stress (MS). Experiments were conducted at the Illinois Institute of Technology under an Institutional Review Board (IRB)-approved protocol. The types of APS inducements were determined based on a literature review of APS inducement research and consultations with a psychologist at the University of Illinois at Chicago. SS activities were separated into three categories based on different APS inducements: SS–non-stress (SS-NS) experiments involved subjects watching a video, resting at home, laying down without APS inducement, and reading a book; SS–mental stress (SS-MS) experiments consisted of multiplications of two-digit numbers under time constraints [11,39,40], solving puzzles, taking IQ or Stroop tests [11,41,42], and playing chess; SS–exciting-anxiety stress (SS-EAS) experiments included watching thriller/horror movies [11,43,44], taking a class exam that did not involve mathematical calculations (to reduce MS) [45], participating in research meetings with advisor, making a presentation to an audience, and playing a video game while sitting [44,46]. TR experiments were conducted on a treadmill while running at a speed in the range of 2.5 to 7.0 mph. The speed was determined by the desired speed of the subject. The TR activities were separated into three categories based on different APS inducements: TR-NS experiments were conducted during TR exercise while listening to calming music or watching natural videos; TR-EAS experiments consisted of watching surgery or car-crash videos during TR exercise; TR-MS experiments were conducted by asking for the mental multiplication of two-digit numbers under time constraints

while running on the treadmill [11,39,40]. APS inducements during stationary bike exercise were similar to implementations in the TR-APS protocol. Information on the experiments is presented in Table 1 and Figure 3.

Table 1. Experiments conducted for physical activity (PA) and acute psychological stress (APS) data collection.

Different Physical State with Various APS Inducements				
Physical Activity	Number of Experiments	Number of Subjects	Duration (min)	Duration After Class Balancing (min)
Resting	89	10	3172	2141
Treadmill	57	20	2164	2141
Stationary Bike	61	19	1713	2141
Sedentary State Experiments with APS Inducement				
APS Type	Number of Experiments	Number of Subjects	Duration (min)	Duration After Class Balancing (min)
Non-Stress	28	6	846	1057
Excitement	29	9	1129	1057
Mental Stress	32	6	1197	1057
Treadmill Experiments with APS Inducement				
APS Type	Number of Experiments	Number of Subjects	Duration (min)	Duration After Class Balancing (min)
Non-Stress	28	20	1162	407
Excitement	12	12	676	407
Mental Stress	17	8	326	407
Stationary Bike Experiments with APS Inducement				
APS Type	Number of Experiments	Number of Subjects	Duration (min)	Duration After Class Balancing (min)
Non-Stress	29	19	891	296
Excitement	24	12	585	296
Mental Stress	8	7	237	296

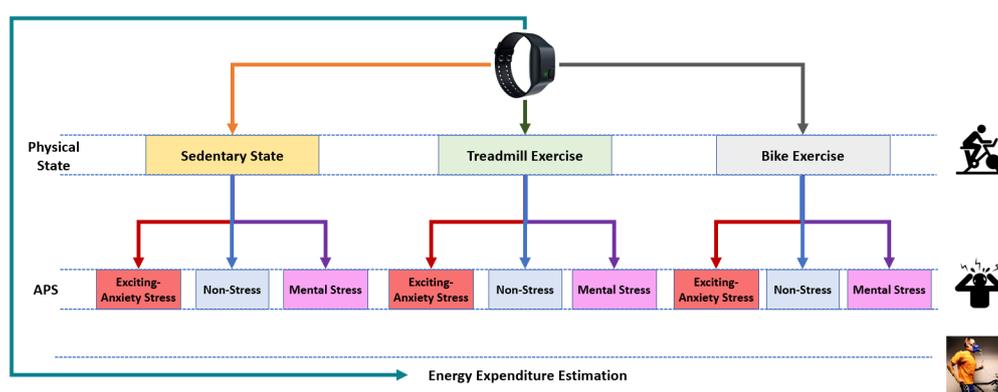


Figure 3. Overview of physiological and psychological state estimation using wristband biosignals (two-layer tree classification approach. 1st layer: Physical state (PS) classification; 2nd layer: APS classification for each PS).

EE is the total amount of energy an individual uses to maintain essential body functions and the energy expended as a result of PA. The gold standard measurement of EE is indirect calorimeter. However, collecting these measurements is uncomfortable, inexpensive, and practical for daily life usage. Previous studies showed that EE can be estimated with wearable devices using biosignals such

as heart rate. However, estimating EE is a still challenging task, though it is a valuable metric for quantifying the intensity of physical activities. We collected data during TR and SB exercises to develop models that can predict the EE using only wristband biosignals. A subset of the experiments for APS and PS model development was used for collecting data for EE estimation. During these experiments, an indirect calorimeter (COSMED K5 wearable metabolic measurement system, COSMED Srl, Italy [47]) data were collected in addition to wristband data. The indirect calorimeter is a gold standard EE measurement that was used as the target output value for training the ML algorithms with the inputs as the wristband data. Sixteen hours of data were collected from 59 experiments with 15 subjects (Table 2).

Table 2. Experiments conducted for energy expenditure (EE) data collection.

Data Collection for EE Model Development (VO2 Mask + Wristband)			
Physical Activity	Number of Experiments	Number of Subjects	Minutes
Treadmill	22	6	304
Stationary Bike	37	9	656

2.2. Signal Processing

In this work, a single non-invasive wrist-worn device, the Empatica E4 wristband (Empatica E4, Empatica Inc., Cambridge, MA, USA [48]) was used. The E4 is capable of recording and streaming the photoplethysmogram (PPG) signal, reporting blood volume pulse (BVP) at a frequency of 64 samples/s, data from a triple-axis accelerometer (ACC) measured at a frequency of 32 samples/s, heart rate derived from the BVP signal using a propriety algorithm and reported at a frequency of 1 samples/s, and ST and GSR recorded at a frequency of 4 samples/s. In this section, signal processing methods are introduced for the five different biosignals (ACC, HR, BVP, GSR, and ST).

2.2.1. Blood Volume Pulse

Since the PPG is highly susceptible to noise and motion artifacts, the PPG signal requires denoising and the ACC data are used as a reference signal to remove the motion artifacts. We used three different signal processing techniques to obtain the denoised PPG signal from the raw PPG data (Figure 4).

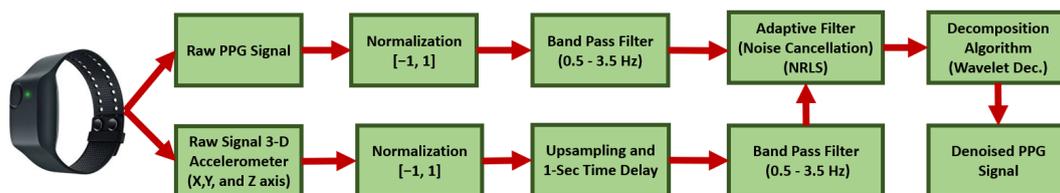


Figure 4. Overview of photoplethysmogram (PPG) signal processing.

Both the raw ACC and PPG signals were normalized within the $[-1,1]$ range. The ACC data were upsampled to obtain the same sampling frequency with the PPG data. Cross-correlation analysis was used to determine the delay between the ACC and the PPG signals. A time-delay of approximately 1 s was found between the ACC and the PPG data, where the highest correlation value occurred at a lag of 73 samples, and with a sampling frequency of the PPG and upsampled ACC signals of 64 Hz, the time-delay was found to be approximately 1.1 s. The same band-pass filter was applied to both the ACC and PPG readings. Physiologically, the HR in humans varies between 30 and 210 bpm (0.5–3.5 Hz) except for extraordinary situations. Several publications consider the 0.5–3.5 Hz frequencies as cut-off frequencies of a band-pass filter [12,49], which was also used for our band-pass filter design. After considering several different methods such as Butterworth, Chebyshev Type I/II, and Elliptic filter, a 4th-order Butterworth band-pass filter was designed with the selected cut-off frequencies. The band-pass filter yielded a smoother signal by eliminating most of the noise. However,

motion artifacts can still exist after the band-pass filter because its frequency can be inside the pass-band region (0.5–3.5 Hz). Therefore, additional steps were considered to filter out the artifacts from the signal.

Adaptive noise cancellation (ANC) is a commonly used approach for PPG signal processing [50,51]. Several different algorithms have been proposed for ANC, including the recursive least-squares filter and the least-mean-squares filter. Nonlinear recursive least-squares (NRLS) filters were used in this work for ANC. We used ANC involving NRLS filtering to remove the motion artifacts in the BVP signal. Since the motion artifacts are related to the movement of the users, and hence the ACC signals, we use the ACC data as the reference signal in the ANC algorithm. In the proposed approach, the ACC readings and the BVP signals are first processed through a band-pass filter to remove frequencies that are not physiologically reasonable. This band-pass filter removes all frequencies that are not representative of the underlying heart rate variations. After only the physiologically meaningful range of frequencies are retained in both the accelerometer and blood volume pulse signals, NRLS is used to remove the motion artifacts in the BVP signal. Since motion artifacts may arise due to movements in different dimensions, we use the three-axis ACC data. We sequentially remove the motion artifacts that are associated with the x-, y-, and z-axis of the ACC signals though the NRLS algorithm applied in series. The NRLS algorithm [51] consists of Volterra series expansion, which provides additional non-linear terms to handle the non-linearity of motion artifacts. The parameters include the filter length (M), forgetting factor (λ), covariance matrix (P), and weight matrix (w), which are specified as 6, 0.999, $1000 \times I$, and 0, respectively. Figure 5 illustrates the structure of ANC using NRLS. The PPG signal is highly correlated with the ACC readings when motion artifacts are present, and the ANC is applied three times in series for the x, y, and z axes, resulting in the signal with significantly reduced motion artifacts.

Equations (1)–(5) describe the mathematical representation of adaptive filter implementation, with $e(k)$ as an enhancement signal; U as the accelerometer measurements; $s(k)$ as the ideal signal, which does not contain any noise; $d(k)$ as motion artifacts or noise desired to be removed; and $d'(k)$ as the noise estimates from adaptive filter. The adaptive filter tries to minimize $d(k) - d'(k)$ using Equations (1)–(5), to obtain a signal as close as possible to $s(k)$, and the filtering equations are as follows:

$$e(k) = s(k) + (d(k) - d'(k)) \quad (1)$$

$$\alpha(k) = \frac{\lambda^{-1}P(k-1)U(k)}{1 + \lambda^{-1}U^T(k)P(k-1)U(k)} \quad (2)$$

$$\omega(k) = \omega(k-1) + \alpha(k)e(k) \quad (3)$$

$$P(k) = \lambda^{-1} \left(1 - \alpha(k)U^T(k) \right) P(k-1), P(0) = \gamma I_M \quad (4)$$

$$d'(k) = \omega(k)^T U(k) \quad (5)$$

The detailed diagram for noise cancellation with an adaptive filter is presented in Figure 5.

ANC presents a useful solution for PPG signal processing. However, ANC only provides good results when ACC readings correlate with the PPG signal, which occurs generally during motion artifacts caused by physical activity. Thus, some additional noise may still remain after ANC. Decomposition algorithms (wavelet decomposition, empirical mode decomposition, and singular signal decomposition) are able to break down signals into different constituent signals at different sub-frequencies. Some sub-frequency sets may be related to additional noise, which can be easily eliminated from the signal. Wavelet decomposition was used in this work for further removing noise from the PPG signal processing [12,51,52]. In this work, the Symlets 4 wavelet function was used with 4 decomposition levels for denoising the PPG signal.

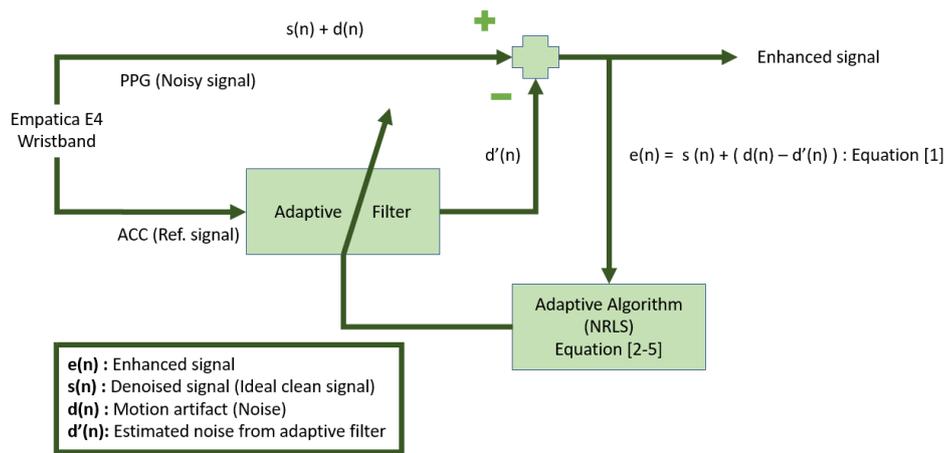


Figure 5. Noise cancellation and artifact removal with nonlinear recursive least-squares (NRLS) adaptive filter.

The efficacy of the ANC and denoising algorithms was evaluated using a ground-truth measurement. A few experiments were conducted with the bioPLUX (bioPLUX, PLUX Wireless Biosignals SA, Portugal [53]), which is a finger-worn device. The bioPLUX collects PPG data at a 1 kHz frequency compared to the 32 Hz frequency of the Empatica E4. The bioPLUX is worn on a hand held constant to reduce motion artifacts and interference, and the data from the bioPLUX were used as a ground-truth to evaluate the improvement made by the PPG filtering and artifact removal algorithm. Data were collected from the other wrist with Empatica E4 during TR without any restrictions in movement. The results of the signal processing algorithms were compared with raw Empatica E4 BVP measurements, the Empatica E4 BVP data processed with the proposed technique, and the bioPLUX measurements as the ground-truth data. Figure 6 presents an example of the improvement in the raw BVP signal achieved with the proposed method. The time frame (5 s) was scaled with common up-sampled frequency (64 kHz) and data were normalized to the [0–1] range. The processed signal significantly enhanced the signal quality.

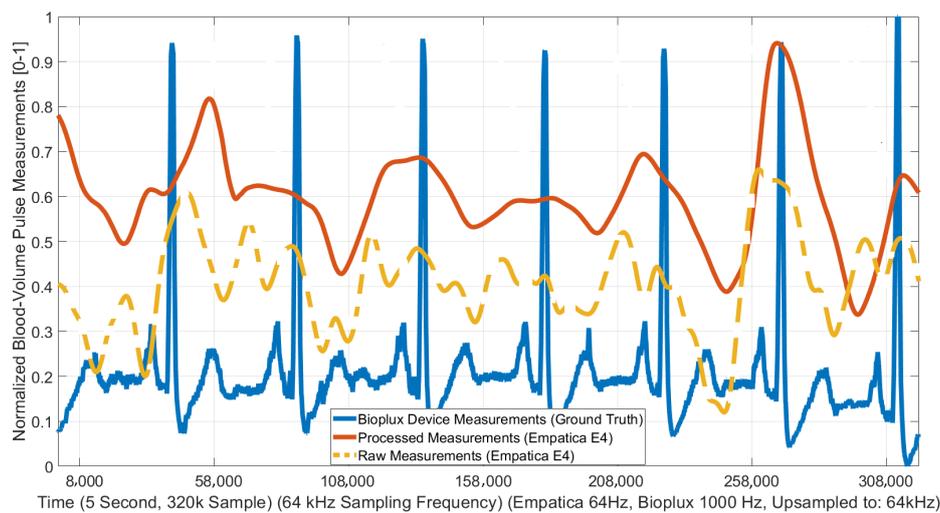


Figure 6. Blood-volume pulse (BVP) signal processing improvements after signal processing implementations (During treadmill (TR) exercise).

2.2.2. Other Biosignals

The Savitzky–Golay (SG) filter was used for denoising the ACC [54], ST [55], HR [11], and GSR [55] measurements, whereas the BVP signal was processed to remove artifacts and noise. The process of

tuning the filtering parameters involves iterative trial-and-error while maximizing the classification accuracy. Different values for the order and frame length were evaluated to find the best filter parameters based on the classification outcomes. After tuning the parameters of the SG filter, the order and frame length parameters were determined for the measurements as: ACC (order = 7 and frame length = 15), ST (5 and 9), HR (5 and 9), and GSR (5 and 11). Signal-to-noise ratio (SNR) is a measure that compares the level of a desired signal to the level of background noise. SNR is defined as the ratio of signal power to the noise power, often expressed in decibels. Higher numbers generally mean a better specification, since there is more useful information (the signal) than unwanted data (the noise). The SNR is calculated as:

$$SNR = P_{\text{signal}}/P_{\text{noise}}, \quad (6)$$

where P_{signal} denotes the average power of the signal and P_{noise} denotes the average power of the noise. We calculated the SNR values for each signal before and after the specified signal processing algorithms. A randomly selected data set was analyzed through the filtering algorithms and the improvements were assessed. The findings showed that SNR of the processed signals were consistently higher than raw measurements (overall SNR values improved 57% (absolute percentage changes)), which showed that the signal cleaning algorithms improve the signal quality by reducing the noise in the signals.

2.3. Feature Extraction and Dimensionality Reduction

A total of 866 primary features were extracted from the five biosignals measured by the wristband device. The features were generated from each one-minute epoch of data. We generated statistical and mathematical features for each variable, including mean, standard deviation, median, quartiles, mean-of-squared values, maximum, minimum, range, slope, first derivative, second derivative, ratio of maximum to minimum values, sum of absolute values, mean of absolute values, interquartile range, coefficient of variation, autoregressive parameters, wavelet decomposition coefficients, signal-to-noise ratio (SNR), skewness, kurtosis, correlations between biosignals, mean normalized frequency, power, magnitude of frequency response, and Fourier transform (Figure 7 and (Table 3)) [11]. These features constitute the primary features and were computed for each variable, yielding 866 primary features. The secondary features were computed as the ratio of primary features from two different biosignals, which yielded 1350 secondary features. The illustration of 10 different commonly-used extracted features from GSR measurements is presented Figure 7. Since the denoised BVP signal significantly eliminates motion artifacts and provides peaks that are accurate (Figure 6), the denoised BVP values were used in feature extraction. Secondary features were obtained by the ratio of primary features to each other, which can indicate some physiological and psychological states. For example, the ratio of ACC to GSR can be useful in discriminating APS from PA because GSR increases during APS while ACC measurements maintains the same level. Finally, 2216 different features were obtained by combining the primary and secondary features. Features were normalized to give each feature variable an equal weight, and principal component analysis (PCA) and partial least squares (PLS) were conducted for feature dimension reduction [56].

The same pool of feature variables was used for PS classification and EE estimation (Table 4). Most of the extracted features are highly correlated with each other (such as median of HR and mean of HR). Highly correlated features do not provide additional information for ML algorithms, and their use may add bias to similar attributes, leading to poor estimation and classification accuracy. PCA is used to reduce the dimensionality of extracted features [57]. This facilitates reducing the number of variables used for training the models, resulting in principal components being retained that capture the informative variation from the feature variables to avoid the risk of overfitting. The first 275 principal components were used as the inputs of each ML algorithm. We varied the number of principal components that were retained when building the models while observing the accuracy of a validation data set. Our analysis showed that retaining 275 principal components resulted in the highest accuracy. Increasing the number of principal components beyond 275 resulted in no improvement in classification accuracy, but the computational requirements increased due to the larger

number of inputs to the models. For EE estimation, PLS regression was used to reduce the features to a smaller set of latent variables that are most correlated with the EE.

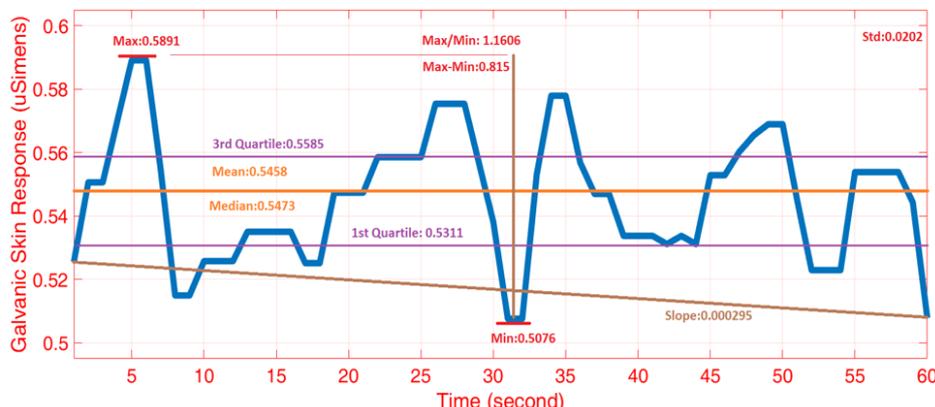


Figure 7. Example feature extraction for galvanic skin response (GSR) (illustration of common statistical and mathematical features such as minimum, quartiles, mean, and slope).

Table 3. List of extracted primary features (BVP-Pr: Processed BVP signal). (*ACC: Magnitude of ACC-X, ACC-Y, and ACC-Z; **ACC: Total number of extracted features from ACC-X, ACC-Y, ACC-Z, and ACC).

Biosignal	Feature	Number of Features
ACC-X	Statistical Features (SD, mean, median, quarterlies, ...)	** ACC: 225
ACC-Y	Mathematical Features (Slope, derivatives related, max-min,	BVP-Raw: 148
ACC-Z	max/min, absolute summation, absolute inverse cosine, zero-cross,	BVP-Pr: 148
* ACC	difference/ratio of 1st and 3rd quarterlies, mean/SD, mean/median,	HR: 98
BVP-Raw	autoregressive all-pole model parameters, logarithm related features,	ST: 111
BVP-Pr.	wavelet decomposition coefficients, signal-to-noise ratio (SNR),...	GSR: 136
HR	Moment Features (Central (different orders), kurtosis, skewness)	
ST	Frequency Domain (Amplitude and location of peak frequency,	
GSR	mean normalized frequency, power, magnitude, FFT features,...)	
	Correlations between other biosignals (Coefficients of	
	(ACC-X, GSR),...), (ACC-X, ACC-Y), (ST, HR), (ACC-X, GSR), ...)	
	Ratio of some fundamental features (SD ACC-X / SD ACC-Z),...	Total: 866
BVP	Heart Rate Variability Features (low frequency, high frequency,	
	very low frequency, ultra low frequency, LF/HF, number of peaks,...)	
GSR	Data Specific Features (mean energy, variance of local peaks,	
ST	total energy response, mean amplitude of local peaks,	
	low frequency, high frequency, very low frequency, LF/HF)	

Due to the lack of physical exertion, the SS experiments were longer in duration than the other experiments, which caused an imbalance among the classes. Since imbalances in class sizes may result in bias and poor classification accuracy, a combination of up-sampling (adaptive synthetic sampling (ADASYN) [58]) and down-sampling methods were used. The impact of upsampling the minor class is reduced by simultaneously downsampling the major class by retaining the unique samples as determined by the similarity measure of k-means clustering. We also studied the effect of different levels of upsampling the minor class on the accuracy of the physical state and acute psychological stress classification results. We found that the highest accuracy was achieved with 25% upsampling of the minor class. After balancing the data set, 2141 min of training data were obtained for each of the SS, TR, and SB activities. In addition, 1057 min of data were obtained for each of NS, EAS, and MS experiments during the SS; 407 min of data were obtained for each of NS, EAS, and MS experiments during SB exercise; and 296 min of data were obtained for NS, EAS, and MS experiments during TR exercise. The final data set sizes of the balanced class sizes after upsampling and downsampling are reported in Table 1 (Column 5).

Table 4. Example of dominant features.

Feature Name	<i>p</i> -Value	Classification/Estimation
SD of ACC/Mean of ACC	≤0.01	PS Classification
Maximum of ST/Maximum of GSR	≤0.01	APS-SS Classification
Mean of HR	≤0.01	APS-TR Classification
Total Energy Response of GSR	≤0.01	APS-SB Classification
Zero cross of BVP	0.49 (correlation)	EE Estimation

2.4. Machine Learning Algorithms

We normalized the biosignals before generating the features to ensure that the absolute magnitudes of some variables did not bias the feature generation. After the features were generated, we normalized the feature variables to ensure each feature variable had equivalent weight in the principal component analysis for dimension reduction. Five different ML models were developed using the normalized latent variables, including the PS classification model, the APS classification model during SS activities, the APS classification model during TR activities, the APS classification model during SB activities, and the EE estimation model. Data were randomized and separated into a training (85%) and testing (15%) data set. The training data set is used for model development and hyperparameter optimization. The testing data set is used to report the final ML outcomes. Various ML algorithms were tested for classification, including, naïve Bayes (NB), k-nearest neighbor (k-NN), decision tree (DT), ensemble learning (EL), support vector machine (SVM), and linear discriminant (LD). Additionally, Gaussian process regression (GPR) was used for EE estimation model development. The k-NN, EL, and SVM models achieve better accuracy relative to the other models. The use of neural network (NN) or deep learning (DL) techniques is an active research area, though these techniques typically require a sufficient number of training samples to appropriately determine the model parameters. Therefore, the performance of NN/DL approaches is limited for the relatively smaller data sets employed in biomedical and health research. The algorithms used in this work are introduced briefly in the following subsections.

2.4.1. k-Nearest Neighbor

We used hyperparameter optimization to determine the distance [59], which is the cosine distance metric, to compare relative distances between the feature variables of the testing data with those of the training data. Based on optimization results, 10 neighbors with the shortest sorted distances were identified as being representative of each class. We assigned the class label to the new data according to the most common class within the 10 closest training samples.

2.4.2. Support Vector Machine/Regression

SVM has the advantage of easy adaptation to feature-based classification approaches for feature variables with nonlinear relations [60]. SVM functions are defined by determining a separating hyperplane in the high-dimensional feature space that best distinguishes two classes during the training stage. The determined separating hyperplane is used online with testing data to assign a class to the test data sample based on the feature variable inputs [60]. We used both SVM and SVR to determine PS and estimate EE [60].

2.4.3. Ensemble Learning

EL uses trained multiple learners—in this case, multiple DT—to achieve higher accuracy than the individual models can achieve alone. Different methods are used to establish an EL model, including the boosting and bagging method [61]. Hyperparameters optimization helps to select the best method with the optimum number of learning cycles, learning rate, and minimum leaf size. We developed an EL model to both classify PS and estimate EE [61].

2.4.4. Linear Discrimination

LD uses features to create a linear model. LD finds a linear combination of features that characterizes or separates two or more classes of objects or events. It is unable to capture nonlinearity. Physiological responses may contain some nonlinearities that cannot be handled by a linear model. However, LD shows better performance in small data sets and with relatively small numbers of iterations for hyperparameter optimization training.

2.4.5. Gaussian Process Regression

Gaussian process regression (GPR) is a stochastic statistical technique that seeks a multivariate normal distribution within random variables from the feature space domain. GPR measures the similarity between points (the kernel function) to predict the value from training data.

2.4.6. Other ML Algorithms

DT and NB are other popular ML algorithms that did not perform well in this work. NB models are based on statistical information and compute the variance and mean of the selected subset of feature variables for a specific number of clusters. We computed the Gaussian and kernel probability density for each feature and computed posterior probability values for the clusters. The computed posterior values were compared to determine the classification. DT generates a predictive model that can be used as a classification or regression model. DT algorithms are highly affected by variations in data, which causes a lack of robustness.

2.4.7. Hyperparameter Optimization

The various hyperparameters for the different ML techniques were optimized with Bayesian optimization (expected improvement acquisition function) to achieve the best classification accuracy [62]. Table 5 lists the hyperparameters for each ML algorithm. A large number of iterations were performed for each ML algorithm using 10-fold cross-validation techniques.

Table 5. List of optimized hyperparameters for the machine learning (ML) algorithms (PS: physical state classification; APS: acute psychological stress; SS-APS: APS classification in sedentary state; TR-APS: APS classification in treadmill run; SB-APS: APS classification on stationary bike; SVM: support vector machine; DT: decision tree; k-NN: k-nearest neighbor; LD: linear discriminant; NB: naive Bayes; EL: ensemble learning)

ML Algorithm	Hyperparameter (s)	PS	SS-APS	TR-APS	SB-APS
SVM	Kernel Scale	0.90	0.46	30.11	0.01
	Box-Constraint	0.19	0.090	122.38	737.2
DT	Minimum Leaf Size	1	7	9	8
k-NN	Number of Neighbors	1	10	7	2
	Distance Function	cosine	Spearman	correlation	correlation
LD	Delta	3.86×10^{-5}	2.1×10^{-5}	0.01	0.0002
	Gamma	0.005	0.0662	0.51	0.22
NB	Distribution Name	normal	normal	kernel	normal
	Width	—	—	0.0472	—
EL	Learning Rate	0.9027	0.001	0.82	0.9504
	Number of Learner	490	16	222	477

3. Results

3.1. Physical State Classification and Psychological Stress Detection

The classification accuracy was calculated with following formula, where true positives (TPs), true negatives (TNs), false positives (FPs), and false negatives (FNs) define the accuracy as:

$$Accuracy(\%) = (TP + TN) / (TP + TN + FP + FN). \tag{7}$$

The various ML algorithms were tested for PS classification. The EL algorithm achieved the best classification accuracy at 99.3% (Figure 8a). The accuracies reported are based on 963 min (approximately 16 h) of testing data (15% of the whole data). The results are presented with a confusion matrix (Figure 8a). Various ML algorithms, SVM, LD, DT, NB, and k-NN, yielded accuracies of 99.0%, 99.7%, 96.6%, 95.6%, and 89.5%, respectively.

The ML algorithms were tested for APS classification (NS, EAS, and MS) during SS activities. The LD algorithm achieved the best accuracy with 98.3% (Figure 8c). The accuracies reported are based on of 473 min (approximately 8 h) of testing data (15% of the whole data), and are presented using a confusion matrix (Figure 8c). Various ML algorithms, SVM, LD, NB, DT, and k-NN, yielded accuracies of 97.6%, 97.7%, 94.0%, 92.6%, and 88.3%, respectively.

The same ML algorithms were also tested for APS classification (NS, EAS, and MS) during TR exercise. The LD algorithm achieved the highest accuracy with yielding 96.7% accuracy (Figure 8b). The accuracies reported are based on 180 min (approximately 3 h) of testing data (15% of the whole data) and are presented in a confusion matrix (Figure 8b). Various ML algorithms, k-NN, EL, SVM, NB, and DT, yielded accuracies of 95.5%, 94.4%, 89.4%, 88.8%, and 84.4%, respectively.

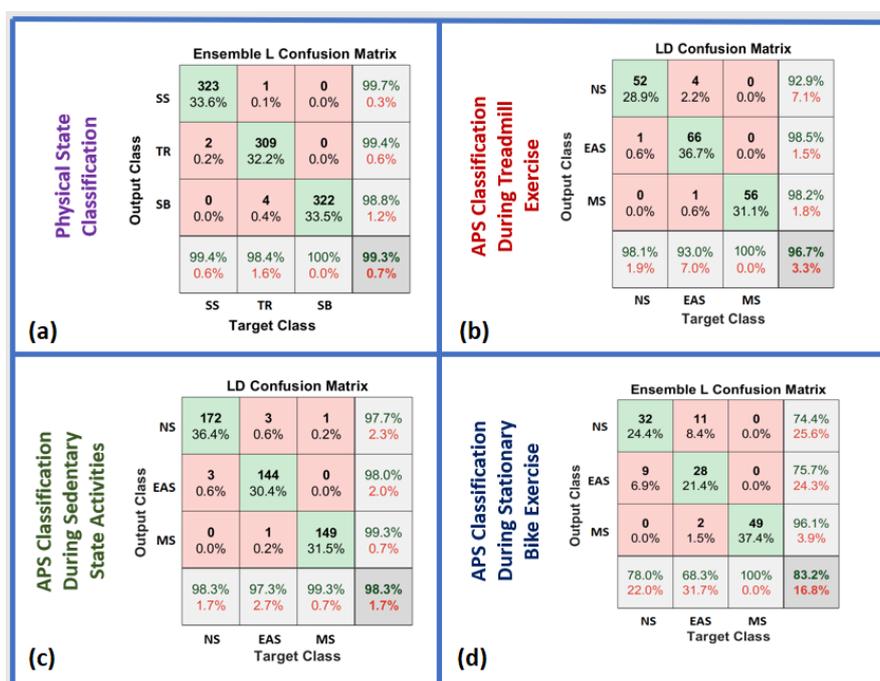


Figure 8. Confusion matrix for PS and APS classification: (a) PS classification with SVM, (b) APS classification during TR exercise with SVM, (c) APS classification during SS activities with EL, and (d) APS classification during SB exercise with k-NN.

The same ML algorithms were also tested for APS classification (NS, EAS, and MS) during SB exercise. The EL algorithm achieved the highest accuracy at 83.2% (Figure 8d). The accuracies reported are based on of 131 min (approximately 2 h) of testing data (15% of the whole data) and are presented

in a confusion matrix (Figure 8d). Various ML algorithms, LD, SVM, DT, NB, and k-NN, yielded accuracies of 81.6%, 80.9%, 71.6%, 70.2%, and 60.3%, respectively.

3.2. Energy Expenditure Estimation

Estimation error was evaluated using mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE), which are computed as follows:

$$MAE(MET) = \frac{1}{n} * \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (8)$$

$$RMSE(MET) = \sqrt{\frac{1}{n} * \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (9)$$

$$MAPE(\%) = \frac{100\%}{n} * \sum_{i=1}^n \left| \frac{(y_i - \hat{y}_i)}{y_i} \right|, \quad (10)$$

where y_i and \hat{y}_i for $i := \{1, \dots, n\}$ denote the actual and predicted values of the EE, respectively.

EE estimation performance was evaluated with 144 min of testing data (15% of the whole data). The estimated values were compared with the indirect calorimeter data. A very low estimation error was achieved with the k-NN algorithm: 0.59% MAPE, 0.02 metabolic equivalent of task (MET) MAE, and 0.06 MET RMSE. Figure 9 illustrates testing data set estimation versus indirect calorimeter measurements. The different algorithms performed significantly differently. The Gaussian process algorithm performed significantly worse with the following estimation error: 29.3% MAPE, 1.66 MET MAE, and 2.09 MET RMSE.

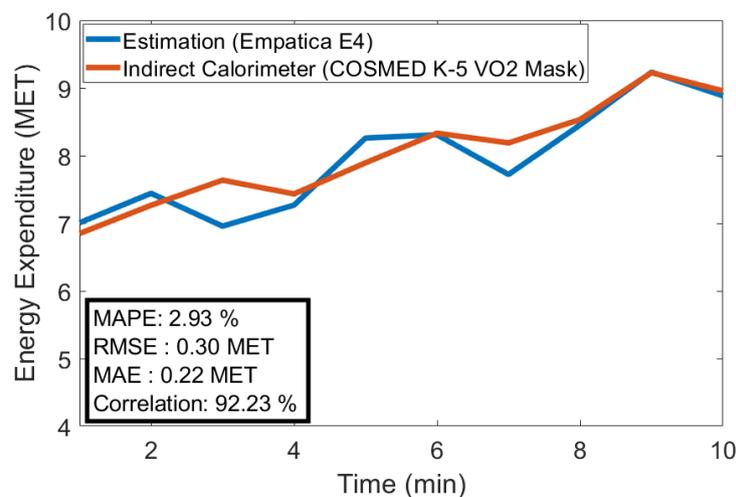


Figure 9. EE estimation (separate data set).

4. Discussion

This paper presents an application for obtaining psychological and physiological states from the signals of a wearable device that can be used in everyday life—a wristband. The performance of several signal processing algorithms and ML methods were assessed. k-NN, NB, and DT performed worse than the other algorithms because these algorithms require more data. However, SVM, LD, and EL produced more accurate performance since they can work with a relatively small amount data with a low risk of overfitting.

The GPR is a powerful ML model; however, it requires a large amount of data. For this reason, in our EE estimation model, it had a large estimation error. SVM, EL, and k-NN achieved similar results that were significantly better than those of the other ML techniques considered.

The signal processing algorithm denoised the wristband PPG signal and removed the artifacts to clearly show the peaks, enabling an accurate estimation of HR. Signals from wristband devices are corrupted with artifacts caused by arm and wrist motion during most physical activities. The developed algorithm improved HR estimation during PA (Figure 6). The contributions of signal processing to accuracy enhancements were evaluated by comparing processed and raw biosignals for the same implementation of various ML algorithms used in this work. Signal processing improved classification accuracies by 0.7% to 4.5%. It also significantly improved the EE estimation. Table 6 presents the improvements with the use of filtered data.

During SS activities, NS was distinguished from other types of APS with a high accuracy (97.7%; Figure 8). However, during exercise, NS could not be discriminated as accurately from other types of APS (92.9–74.4%; Figure 8). The accuracy decreased because exercise and APS can affect some biosignals such as HR. Similarly, both APS and exercise increase HR.

Filtering improves signal quality by smoothing the signals and reducing the amount of noise, sharp and sudden changes, and outliers. Features are extracted from both filtered and raw data. Features with filtered data represent reasonable distributions in a limited range scale with fewer outliers, which is crucial because features are used for the design of all ML algorithms. The raw BVP signal is also used for feature extraction because motion artifacts can help capture the type of PA and discriminate PA from APS. Figure 10 illustrates the improvements with filtering for various biosignals. Signal processing algorithms also improve the classification and estimation outcomes. Filtered signals performed better by up to 3% for APS classification and 18% for EE estimation (Table 6).

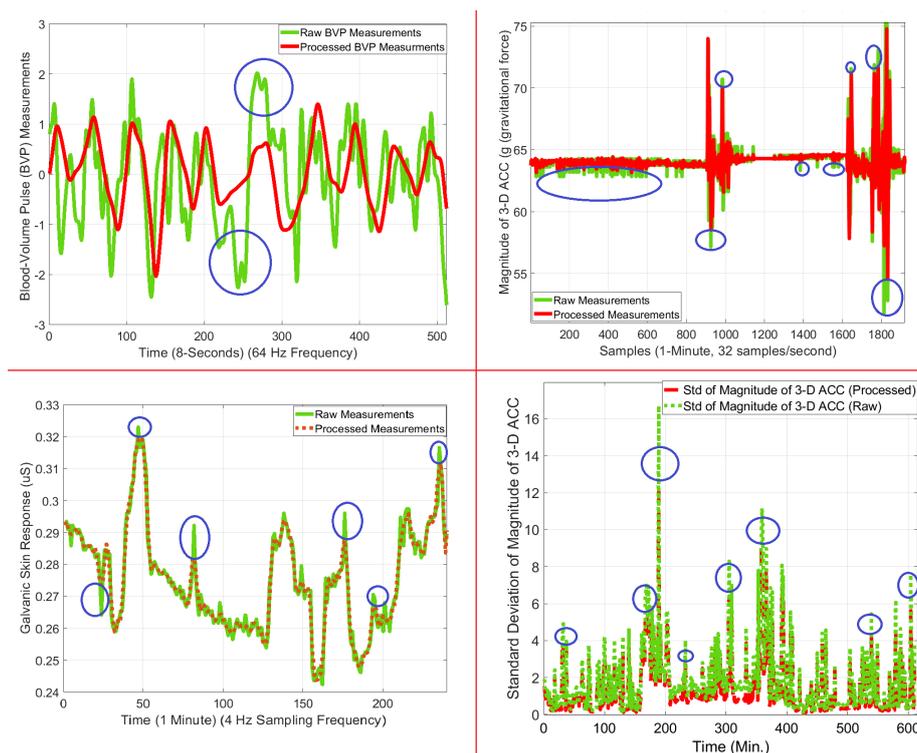


Figure 10. Filtering and signal processing improvements for raw biosignals: (**Upper Left**) BVP signal processing improvements during TR exercise, (**Upper Right**) ST data filtering result during SB exercise, (**Lower Left**) GSR data filtering result SB exercise, and (**Lower Right**) feature extraction example comparison with and without signal processing for ACC data during SS activities. (Blue circles: improvements where there are visible abrupt changes).

Table 6. Signal processing effects on accuracy improvements. MAE: mean absolute error; RMSE: root-mean squared error; MAPE: mean absolute percentage error.

ML Algorithm	Raw Biosignals Best Testing Accuracy (%)	Processed Biosignals Best Testing Accuracy (%)
PS Classification	98.6	99.3
APS Classification SS	96.2	98.3
APS Classification TR	93.3	96.6
APS Classification SB	79.4	83.2

ML Algorithm	Raw Biosignals Estimation Error	Processed Biosignals Estimation Error
EE Estimation (MAPE %)	0.72	0.59
EE Estimation (MAE (MET))	0.04	0.02
EE Estimation (RMSE (MET))	0.13	0.06

We analyzed the effect of each sensor modality on the classification accuracy. In this analysis, we excluded the feature variables derived from ACC, GSR, HR, BVP, and ST measurements, and built the classification models with the retained measurements only, which included the PS, APS-SS, APS-TR, and APS-SB classifications. The results showed a decrease in classification accuracy when any individual measurement variable was removed, thus demonstrating the advantage of the multisensor fusion method in improving the classification accuracy (Figure 11). The different modalities of sensing physiological variables provide various contributions to the classification accuracy. For example, the galvanic skin response measurements contribute significantly to improving the psychological stress during biking relative to the other biosignals. The accelerometer signals contribute to improving the physical state classification.

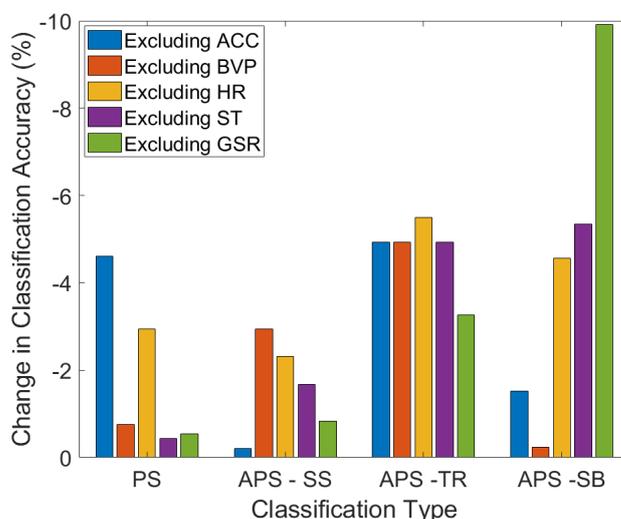


Figure 11. Relative (all biosignals) classification accuracy changes with excluded biosignals for each classification type.

One limitation of the current work is that the data collected were not sufficient to develop advanced deep learning models; because of this limitation, we found that other machine learning techniques performed better than the deep learning approaches. More data are needed to appropriately train the advanced deep learning models and these results should not be considered for assessing deep learning models.

Our research focused on identifying the effects of physical activity and acute psychological stress on the glucose–insulin dynamics of people with Type 1 diabetes. Due to insufficient insulin secretion, people with Type 1 diabetes must administer exogenous insulin to regulate their blood glucose levels. The amount of insulin dose to administer depends on the physical and psychological state of the

subject. The algorithms developed in this work will enable real-time assessment of the physical and psychological stressors experienced by people with Type 1 diabetes and how their insulin requirements vary based on these factors.

5. Conclusions

Signal processing and ML algorithms were used on five different biosignals reported by a single wristband to detect, classify, and quantify physical activities and acute psychological stress. The signals processed for noise and artifact elimination were used for feature extraction, feature reduction, and feature use in the development of various ML models. The ML models are able to accurately detect and identify the characteristics of PA and APS and estimate EE. The improvements in the accuracy for detecting and characterizing PA and APS and their concurrent occurrence improves the feasibility of using this information in treatment decisions systems of chronic diseases such as automated insulin dosing decisions in diabetes.

Author Contributions: M.S., M.R., M.R.A., I.H. and A.C. conceived the research. M.S. and M.R. developed the theory and performed the computations. M.S. and Z.M. organized the experiments. M.S., M.R. and A.C. wrote the manuscript. All authors discussed the results and contributed to the final manuscript. A.C. supervised the project. All authors have read and agreed to the published version of the manuscript.

Funding: Financial support is gratefully acknowledged from the NIH under the grants 1DP3DK101075 and 1DP3DK101077 and JDRF under grant 2-SRA-2017-506-M-B, which was made possible through collaboration between the JDRF and The Leona M. and Harry B. Helmsley Charitable Trust.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACC	Accelerometer
ADASYN	Adaptive Synthetic Sampling Approach
APS	Acute Psychological Stress
BVP	Blood-Volume Pulse
DT	Decision Tree
EAS	Exciting-Anxiety Stress
EE	Energy Expenditure
EL	Ensemble Learning
GSR	Galvanic Skin Response
GPR	Gaussian Process Regression
HR	Heart Rate
k-NN	k-Nearest Neighbor
LD	Linear Discrimination
LR	Linear Regression
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MET	Metabolic Equivalent of Task
ML	Machine Learning
MS	Mental-Stress
NB	Naïve Bayes
NS	Non-Stress
PCA	Principal Component Analysis
PLS	Partial Least Squares
PPG	Photoplethysmogram
PS	Physical State
RMSE	Root Mean Square Error
SB	Stationary Bike
SG	Savitzky–Golay
ST	Skin Temperature
SS	Sedentary State
SVM	Support Vector Machine
TR	Treadmill

References

1. Zheng, Y.L.; Ding, X.R.; Poon, C.C.Y.; Lo, B.P.L.; Zhang, H.; Zhou, X.L.; Yang, G.Z.; Zhao, N.; Zhang, Y.T. Unobtrusive sensing and wearable devices for health informatics. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 1538–1554. [[CrossRef](#)] [[PubMed](#)]
2. Coughlin, S.S.; Stewart, J. Use of consumer wearable devices to promote physical activity: A review of health intervention studies. *J. Environ. Health Sci.* **2016**, *2*. [[CrossRef](#)] [[PubMed](#)]
3. Chiauzzi, E.; Rodarte, C.; Das Mahapatra, P. Patient-centered activity monitoring in the self-management of chronic health conditions. *BMC Med.* **2015**, *13*, 1–6. [[CrossRef](#)] [[PubMed](#)]
4. Cheol Jeong, I.; Bychkov, D.; Searson, P.C. Wearable devices for precision medicine and health state monitoring. *IEEE Trans. Biomed. Eng.* **2018**, *66*, 1242–1258. [[CrossRef](#)]
5. Can, Y.S.; Arnrich, B.; Ersoy, C. Stress detection in daily life scenarios using smart phones and wearable sensors: A survey. *J. Biomed. Inform.* **2019**, *92*, 103139. [[CrossRef](#)]
6. Turksoy, K.; Hajizadeh, I.; Hobbs, N.; Kilkus, J.; Littlejohn, E.; Samadi, S.; Feng, J.; Sevil, M.; Lazaro, C.; Ritthaler, J.; et al. Multivariable artificial pancreas for various exercise types and intensities. *Diabetes Technol. Ther.* **2018**, *20*, 662–671. [[CrossRef](#)]
7. Giannakakis, G.; Grigoriadis, D.; Giannakaki, K.; Simantiraki, O.; Roniotis, A.; Tsiknakis, M. Review on psychological stress detection using biosignals. *IEEE Trans. Affect. Comput.* **2019**. [[CrossRef](#)]
8. Yang, C.C.; Hsu, Y.L. A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors* **2010**, *10*, 7772–7788. [[CrossRef](#)]
9. Turksoy, K.; Paulino, T.M.L.; Zaharieva, D.P.; Yavelberg, L.; Jamnik, V.; Riddell, M.C.; Cinar, A. Classification of physical activity: Information to artificial pancreas control systems in real time. *J. Diabetes Sci. Technol.* **2015**, *9*, 1200–1207. [[CrossRef](#)]
10. Sevil, M.; Rashid, M.; Maloney, Z.; Hajizadeh, I.; Samadi, S.; Askari, M.R.; Hobbs, N.; Brandt, R.; Park, M.; Quinn, L.; et al. Determining Physical Activity Characteristics from Wristband Data for Use in Automated Insulin Delivery Systems. *IEEE Sens. J.* **2020**. [[CrossRef](#)]
11. Sevil, M.; Hajizadeh, I.; Samadi, S.; Feng, J.; Lazaro, C.; Frantz, N.; Yu, X.; Brandt, R.; Maloney, Z.; Cinar, A. Social and competition stress detection with wristband physiological signals. In Proceedings of the 2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN), Eindhoven, The Netherlands, 9–12 May 2017; IEEE: New York, NY, USA, 2017; pp. 39–42.
12. Biswas, D.; Simões-Capela, N.; Van Hoof, C.; Van Helleputte, N. Heart rate estimation from wrist-worn photoplethysmography: A review. *IEEE Sens. J.* **2019**, *19*, 6560–6570. [[CrossRef](#)]
13. Sevil, M.; Rashid, M.; Hajizadeh, I.; Maloney, Z.; Samadi, S.; Askari, M.R.; Brandt, R.; Hobbs, N.; Park, M.; Quinn, L.; et al. Assessing the Effects of Stress Response on Glucose Variations. In Proceedings of the 2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN), Chicago, IL, USA, 19–22 May 2019; IEEE: New York, NY, USA, 2019; pp. 1–4.
14. Riddell, M.C.; Zaharieva, D.P.; Yavelberg, L.; Cinar, A.; Jamnik, V.K. Exercise and the development of the artificial pancreas: One of the more difficult series of hurdles. *J. Diabetes Sci. Technol.* **2015**, *9*, 1217–1226. [[CrossRef](#)] [[PubMed](#)]
15. Breton, M.D.; Brown, S.A.; Karvetski, C.H.; Kollar, L.; Topchyan, K.A.; Anderson, S.M.; Kovatchev, B.P. Adding heart rate signal to a control-to-range artificial pancreas system improves the protection against hypoglycemia during exercise in type 1 diabetes. *Diabetes Technol. Ther.* **2014**, *16*, 506–511. [[CrossRef](#)]
16. Kudva, Y.C.; Carter, R.E.; Cobelli, C.; Basu, R.; Basu, A. Closed-loop artificial pancreas systems: Physiological input to enhance next-generation devices. *Diabetes Care* **2014**, *37*, 1184–1190. [[CrossRef](#)] [[PubMed](#)]
17. Wong, J.C.Y.; Wang, J.; Fu, E.Y.; Leong, H.V.; Ngai, G. Activity Recognition and Stress Detection via Wristband. In Proceedings of the 17th International Conference on Advances in Mobile Computing & Multimedia, Munich, Germany, 2–4 December 2019; ACM Digital Library: New York, NY, USA, 2019; pp. 102–106. [[CrossRef](#)]
18. Sun, F.T.; Kuo, C.; Cheng, H.T.; Buthpitiya, S.; Collins, P.; Griss, M. Activity-aware mental stress detection using physiological sensors. In *International Conference on Mobile Computing, Applications, and Services*; Springer: Berlin/Heidelberg, Germany, 2010; pp. 282–301.
19. Salazar-Ramirez, A.; Irigoyen, E.; Martinez, R.; Zalabarria, U. An enhanced fuzzy algorithm based on advanced signal processing for identification of stress. *Neurocomputing* **2018**, *271*, 48–57. [[CrossRef](#)]

20. Indra Indikawati, F.; Winiarti, S. Stress Detection from Multimodal Wearable Sensor Data. *Manag. Sci. Eng.* **2020**, *771*, 12–28. [[CrossRef](#)]
21. Abbas, M.; Jeannès, R.L.B. Characterizing Peaks in Acceleration Signals—Application to Physical Activity Detection using Wearable Sensors. *IEEE Sens. J.* **2020**, *20*, 12384–12395 [[CrossRef](#)]
22. Devasahayam, S.R. *Signals and Systems in Biomedical Engineering: Physiological Systems Modeling and Signal Processing*; Springer: Singapore, 2019. [[CrossRef](#)]
23. Lahmiri, S.; Boukadoum, M. A weighted bio-signal denoising approach using empirical mode decomposition. *Biomed. Eng. Lett.* **2015**, *5*, 131–139. [[CrossRef](#)]
24. Banerjee, R.; Ghose, A.; Choudhury, A.D.; Sinha, A.; Pal, A. Noise cleaning and Gaussian modeling of smart phone photoplethysmogram to improve blood pressure estimation. In Proceedings of the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, Australia, 19–24 April 2015; IEEE: New York, NY, USA, 2015; pp. 967–971.
25. Chong, J.W.; Dao, D.K.; Salehizadeh, S.; McManus, D.D.; Darling, C.E.; Chon, K.H.; Mendelson, Y. Photoplethysmograph signal reconstruction based on a novel hybrid motion artifact detection—Reduction approach. Part I: Motion and noise artifact detection. *Ann. Biomed. Eng.* **2014**, *42*, 2238–2250. [[CrossRef](#)]
26. Askari, M.R.; Rashid, M.; Sevil, M.; Hajizadeh, I.; Brandt, R.; Samadi, S.; Cinar, A. Artifact Removal from Data Generated by Nonlinear Systems: Heart Rate Estimation from Blood Volume Pulse Signal. *Ind. Eng. Chem. Res.* **2019**, *59*, 2318–2327. [[CrossRef](#)]
27. Ziegler, M.G. Psychological stress and the autonomic nervous system. In *Primer on the Autonomic Nervous System*; Elsevier: Amsterdam, The Netherlands, 2012; pp. 291–293.
28. Jarczok, M.N.; Jarczok, M.; Mauss, D.; Koenig, J.; Li, J.; Herr, R.M.; Thayer, J.F. Autonomic nervous system activity and workplace stressors—A systematic review. *Neurosci. Biobehav. Rev.* **2013**, *37*, 1810–1823. [[CrossRef](#)] [[PubMed](#)]
29. Smets, E.; Casale, P.; Großekathöfer, U.; Lamichhane, B.; De Raedt, W.; Bogaerts, K.; Van Diest, I.; Van Hoof, C. Comparison of machine learning techniques for psychophysiological stress detection. In *International Symposium on Pervasive Computing Paradigms for Mental Health*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 13–22.
30. Xia, L.; Malik, A.S.; Subhani, A.R. A physiological signal-based method for early mental-stress detection. *Biomed. Signal Process. Control* **2018**, *46*, 18–32. [[CrossRef](#)]
31. Bornoiu, I.V.; Grigore, O. A study about feature extraction for stress detection. In Proceedings of the 2013 8th International Symposium on Advanced Topics in Electrical Engineering (ATEE), Bucharest, Romania, 23–25 May 2013; IEEE: New York, NY, USA, 2013; pp. 1–4.
32. Ciabattini, L.; Ferracuti, F.; Longhi, S.; Pepa, L.; Romeo, L.; Verdini, F. Real-time mental stress detection based on smartwatch. In Proceedings of the 2017 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 8–10 January 2017; IEEE: New York, NY, USA, 2017; pp. 110–111.
33. Zamil, A.A.A.; Hasan, S.; Baki, S.M.J.; Adam, J.M.; Zaman, I. Emotion detection from speech signals using voting mechanism on classified frames. In Proceedings of the 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), Dhaka, Bangladesh, 10–12 January 2019; IEEE: New York, NY, USA, 2019; pp. 281–285.
34. Steptoe, A.; Kivimäki, M. Stress and cardiovascular disease. *Nat. Rev. Cardiol.* **2012**, *9*, 360–370. [[CrossRef](#)] [[PubMed](#)]
35. Durstine, J.L.; Gordon, B.; Wang, Z.; Luo, X. Chronic disease and the link to physical activity. *J. Sport Health Sci.* **2013**, *2*, 3–11. [[CrossRef](#)]
36. Shrivastava, S.R.; Shrivastava, P.S.; Ramasamy, J. Role of self-care in management of diabetes mellitus. *J. Diabetes Metab. Disord.* **2013**, *12*, 14. [[CrossRef](#)] [[PubMed](#)]
37. Colberg, S.R.; Sigal, R.J.; Yardley, J.E.; Riddell, M.C.; Dunstan, D.W.; Dempsey, P.C.; Horton, E.S.; Castorino, K.; Tate, D.F. Physical activity/exercise and diabetes: A position statement of the American Diabetes Association. *Diabetes Care* **2016**, *39*, 2065–2079. [[CrossRef](#)]
38. Li, L.; Li, X.; Zhou, W.; Messina, J.L. Acute psychological stress results in the rapid development of insulin resistance. *J. Endocrinol.* **2013**, *217*, 175. [[CrossRef](#)]
39. Tanida, M.; Katsuyama, M.; Sakatani, K. Relation between mental stress-induced prefrontal cortex activity and skin conditions: A near-infrared spectroscopy study. *Brain Res.* **2007**, *1184*, 210–216. [[CrossRef](#)]

40. Widjaja, D.; Orini, M.; Vlemincx, E.; Van Huffel, S. Cardiorespiratory dynamic response to mental stress: A multivariate time-frequency analysis. *Comput. Math. Methods Med.* **2013**, *2013*. [CrossRef]
41. Tulen, J.; Moleman, P.; Van Steenis, H.; Boomsma, F. Characterization of stress reactions to the Stroop Color Word Test. *Pharmacol. Biochem. Behav.* **1989**, *32*, 9–15. [CrossRef]
42. Hjemdahl, P.; Freyschuss, U.; Juhlin-Dannfelt, A.; Linde, B. Differentiated sympathetic activation during mental stress evoked by the Stroop test. *Acta Physiol. Scand. Suppl.* **1984**, *527*, 25–29. [PubMed]
43. Palmer, M.A. Fear: A Psychophysiological Study of Horror Film Viewing. Available online: <https://digital.library.txstate.edu/bitstream/handle/10877/3220/fulltext.pdf?sequence=1&isAllowed=y> (accessed on 3 August 2020)
44. Bosse, T.; Gerritsen, C.; de Man, J.; Stam, M. Inducing anxiety through video material. In *International Conference on Human-Computer Interaction*; Springer: Berlin/Heidelberg, Germany, 2014; Volume 434, pp. 301–306.
45. Karthikeyan, P.; Murugappan, M.; Yaacob, S. A review on stress inducement stimuli for assessing human stress using physiological signals. In Proceedings of the 2011 IEEE 7th International Colloquium on Signal Processing and its Applications, Penang, Malaysia, 4–6 March 2011; IEEE: New York, NY, USA, 2011; pp. 420–425.
46. Zhang, Z.; Su, H.; Peng, Q.; Yang, Q.; Cheng, X. Exam anxiety induces significant blood pressure and heart rate increase in college students. *Clin. Exp. Hypertens.* **2011**, *33*, 281–286. [CrossRef] [PubMed]
47. Perez-Suarez, I.; Martin-Rincon, M.; Gonzalez-Henriquez, J.J.; Fezzardi, C.; Perez-Regalado, S.; Galvan-Alvarez, V.; Juan-Habib, J.W.; Morales-Alamo, D.; Calbet, J.A. Accuracy and precision of the COSMED K5 portable analyser. *Front. Physiol.* **2018**, *9*, 1764. [CrossRef]
48. Empatica E4. Available online: <https://www.empatica.com/research/e4/> (accessed on 16 June 2020).
49. Temko, A. Accurate heart rate monitoring during physical exercises using PPG. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 2016–2024. [CrossRef] [PubMed]
50. Chowdhury, S.S.; Hyder, R.; Hafiz, M.S.B.; Haque, M.A. Real-time robust heart rate estimation from wrist-type PPG signals using multiple reference adaptive noise cancellation. *IEEE J. Biomed. Health Inform.* **2016**, *22*, 450–459. [CrossRef] [PubMed]
51. Ye, Y.; Cheng, Y.; He, W.; Hou, M.; Zhang, Z. Combining nonlinear adaptive filtering and signal decomposition for motion artifact removal in wearable photoplethysmography. *IEEE Sens. J.* **2016**, *16*, 7133–7141. [CrossRef]
52. Joseph, G.; Joseph, A.; Titus, G.; Thomas, R.M.; Jose, D. Photoplethysmogram (PPG) signal analysis and wavelet de-noising. In Proceedings of the 2014 Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/iCMMD), Kottayam, India, 24–26 July 2014; IEEE: New York, NY, USA, 2014; pp. 1–5.
53. Bioplux Device. Available online: <https://www.biosignalsplux.com/index.php/bvp-blood-volume-pulse> (accessed on 16 June 2020).
54. Li, Q.; Chen, X.; Xu, W. Noise reduction of accelerometer signal with singular value decomposition and Savitzky-Golay filter. *J. Inf. Comput. Sci.* **2013**, *10*, 4783–4793. [CrossRef]
55. Mohammad, Y.; Nishida, T. Using physiological signals to detect natural interactive behavior. *Appl. Intell.* **2010**, *33*, 79–92. [CrossRef]
56. Subasi, A.; Gursoy, M.I. EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Syst. Appl.* **2010**, *37*, 8659–8666. [CrossRef]
57. Kambhatla, N.; Leen, T.K. Dimension reduction by local principal component analysis. *Neural Comput.* **1997**, *9*, 1493–1516. [CrossRef]
58. He, H.; Bai, Y.; Garcia, E.A.; Li, S. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, China, 1–8 June 2008; IEEE: New York, NY, USA, 2008; pp. 1322–1328.
59. Hwang, W.J.; Wen, K.W. Fast kNN classification algorithm based on partial distance search. *Electron. Lett.* **1998**, *34*, 2062–2063. [CrossRef]
60. Scholkopf, B.; Smola, A.J.; Bach, F. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*; MIT Press: Cambridge, MA, USA, 2018.
61. Amari, S. *The Handbook of Brain Theory and Neural Networks*; MIT Press: Cambridge, MA, USA, 2003.

62. Snoek, J.; Larochelle, H.; Adams, R.P. Practical bayesian optimization of machine learning algorithms. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 2951–2959.

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).